



TECHNISCHE  
UNIVERSITÄT  
WIEN

Vienna University of Technology

## DIPLOMARBEIT

# Implementierung und Vergleich von vier Methoden zur Parameterschätzung bei affinen Modellen mit und ohne Sprünge

Ausgeführt am Institut für  
Stochastik und Wirtschaftsmathematik  
der Technischen Universität Wien

unter der Anleitung von Ao.Univ.Prof. Dr. Friedrich Hubalek

durch

Sabine Sischka

Neudeggasse 17/4, 1080 Wien

---

Datum

---

Unterschrift (Student)

## **Abstract**

Diese Arbeit beschäftigt sich mit Methoden zur Parameterschätzung bei zwei affinen Modellen, dem CIR Modell und dem  $\Gamma$ -OU Modell. Es werden im Speziellen die Maximum-Likelihood Methode, Momentenmethode, exponentielle Momentenmethode und Quasi-Maximum-Likelihood Methode betrachtet, wobei zunächst Schätzer für die jeweiligen drei Parameter erarbeitet werden und schließlich die Varianzen dieser Schätzer untersucht werden. Letztendlich wird zudem die erarbeitete Theorie auf konkrete Beispiele angewendet und veranschaulicht.

**Abstract**

This thesis deals with parameter estimation methods of two affine models, the CIR model and the  $\Gamma$ -OU model. Specifically the maximum likelihood method, the method of moments, the exponential method of moments and the quasi maximum likelihood method will be considered, whereas we will first develop estimators for the three parameters respectively and then analyse the variances of those estimators. Ultimately the theory will be applied to specific examples and visualised.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>6</b>
<b>2</b>	<b>CIR Modell</b>	<b>8</b>
2.1	Bedingte Verteilung . . . . .	9
2.2	Stationäre Verteilung . . . . .	13
2.3	Simulation . . . . .	14
<b>3</b>	<b>Parameterschätzung CIR Modell</b>	<b>16</b>
3.1	Maximum-Likelihood Methode . . . . .	17
3.2	Momentenmethode . . . . .	18
3.3	Exponentielle Momentenmethode . . . . .	21
3.4	Quasi-Maximum-Likelihood Methode . . . . .	23
<b>4</b>	<b>Varianzschätzung CIR Modell</b>	<b>26</b>
4.1	Varianzschätzung Maximum-Likelihood Methode . . . . .	26
4.2	Varianzschätzung Momentenmethode . . . . .	28
4.3	Bemerkung zur Varianzschätzung Exponentielle Momentenmethode . . . . .	31
4.4	Varianzschätzung Quasi-Maximum-Likelihood Methode . . . . .	32
<b>5</b>	<b>Vergleich und numerische Illustrationen CIR Modell</b>	<b>34</b>
<b>6</b>	<b>Gamma-OU Modell</b>	<b>41</b>
6.1	Bedingte Verteilung . . . . .	42
6.2	Stationäre Verteilung . . . . .	47
6.3	Simulation . . . . .	47
<b>7</b>	<b>Parameterschätzung Gamma-OU Modell</b>	<b>50</b>
7.1	Momentenmethode . . . . .	52
7.2	Exponentielle Momentenmethode . . . . .	55
7.3	Quasi-Maximum-Likelihood Methode . . . . .	56
<b>8</b>	<b>Varianzschätzung Gamma-OU Modell</b>	<b>58</b>
8.1	Varianzschätzung Momentenmethode . . . . .	58
8.2	Bemerkung zur Varianzschätzung Exponentielle Momentenmethode . . . . .	61

8.3	Varianzschätzung Quasi-Maximum-Likelihood Methode . . . . .	62
<b>9</b>	<b>Vergleich und numerische Illustrationen Gamma-OU Modell</b>	<b>64</b>

# 1 Einleitung

In dieser Arbeit werden wir verschiedene Schätzmethoden zu zwei in der Finanzmathematik häufig verwendeten affinen Modellen erarbeiten und implementieren.

Zuerst betrachten wir ein stochastisches Volatilitätsmodell, bei dem die Volatilität durch eine Brownsche Bewegung angetrieben wird - das CIR Modell. Insbesondere ist dieses ein nicht-lineares Modell, was notwendig ist, um eine nicht negative Volatilität zu gewährleisten. Während die Brownsche Bewegung der wohl am häufigsten verwendete Prozess in der stochastischen Analysis ist und es nur natürlich scheint ein solches Modell in Betracht zu ziehen, muss man sich doch klar werden, dass die Stetigkeit der Brownschen Bewegung bei der Modellierung von Volatilität beim Vergleich mit echten Daten einen Nachteil darstellt.

Um diesem Problem Abhilfe zu verschaffen, kann man ein Modell betrachten, das Sprünge zulässt. In diesem Sinne betrachten wir als zweites Modell ein Non-Gaussian Ornstein Uhlenbeck (OU) Modell, nämlich das  $\Gamma$ -OU Modell, welches als stochastischen Term einen wachsenden Lévy Prozess besitzt, in unserem genauer studierten Fall ist dieser ein Compound Poisson Prozess.

Wir werden bei beiden Modellen zunächst die Verteilung des Prozesses untersuchen, insbesondere die bedingte Verteilung, die durch die bedingte charakteristische Funktion beschrieben wird. Mit dem Wissen über die Verteilung simulieren wir Daten auf einem diskreten Zeitgitter, wobei zur besseren Vergleichbarkeit der beiden Modelle die Parameter so gewählt werden, dass der stationäre Erwartungswert und die stationäre Varianz gleich sind.

Im zweiten Schritt schätzen wir die Parameterwerte aus unseren simulierten Daten mithilfe verschiedener Methoden. Wir untersuchen dabei vier verschiedene Methoden:

- Die Maximum-Likelihood Methode, die sich der exakten bedingten Übergangsdichten bedient
- Die Momentenmethode, die als Grundlage die bedingten Momente der Verteilung benötigt
- Die exponentielle Momentenmethode, die direkt die bedingte charakteristische Funktion verwendet

- Die Quasi-Maximum-Likelihood Methode, die normalverteilte bedingte Übergangsdichten annimmt

Schließlich machen wir uns Gedanken über die Varianzen der vorhin erarbeiteten Schätzer und geben, wenn möglich, einen Schätzer für die Varianz-Kovarianz Matrix an.

Zuletzt werden die Ergebnisse durch ein konkretes Beispiel veranschaulicht. Neben den theoretischen Varianzschätzern werden zudem numerische Illustrationen in Form von Histogrammen der Schätzer der einzelnen Parameter präsentiert, die durch wiederholtes Anwenden der jeweiligen Schätzmethode entstanden sind.

## 2 CIR Modell

Das CIR Modell, benannt nach John C. Cox, Jonathan E. Ingersoll, Jr. und Stephen A. Ross, ist ein Modell für kurzfristige Zinssätze und wurde erstmals in ihrer Arbeit 'A Theory of the Term Structure of Interest Rates' [5] beschrieben. Die Veränderung der Zinsrate wird hierbei durch die folgende stochastische Differentialgleichung beschrieben:

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t \quad (1)$$

mit Parametern  $\kappa$ ,  $\theta$  und  $\sigma$  und Brownscher Bewegung  $W_t$ .

Wie bekannt, ist eine Brownsche Bewegung ein Zufallsprozess mit unabhängigen, stationären Inkrementen, welche normalverteilt sind.

Die Vorteile einer Zinsrate, die dieser Struktur folgt, sind (siehe [4] und [5]):

- Der Zins wird, sofern der Startwert ein positiver ist, nie negativ und im Falle, dass  $2\kappa\theta \geq \sigma^2$ , bleibt der Zins sogar durchgehend positiv.
- Sollte der Zins Null erreichen, kann er danach wieder positiv werden.
- Der Zinsprozess ist mean-reverting.
- Die Volatilität ist abhängig vom momentanen Level des Prozesses, sie ist also nicht konstant.
- Es gibt eine stationäre Verteilung.

Wie man sieht, verwendet der CIR Prozess als stochastischen Term eine Brownsche Bewegung. Diese ist der wohl am meisten studierte stochastische Prozess und dessen Stetigkeit und Skalierungsinvarianz sind häufig gewünschte Eigenschaften.

Was oft ein Vorteil ist, entspricht bei Vergleich mit Daten nicht immer exakt der Realität. Besonders das Fehlen von Sprüngen ist realitätsfern.

## 2.1 Bedingte Verteilung

Um mit dem CIR Prozess arbeiten zu können, benötigen wir folgendes Wissen über die Verteilung:

**Satz 2.1.** *Mit  $s < t$  und  $\Delta = t - s$  gilt:*

1. *Die bedingte charakteristische Funktion ist*

$$\begin{aligned} \mathbb{E}[\exp(iyr_t)|r_s = r] &= \left(1 - \frac{\sigma^2}{2\kappa}iy(1 - \exp(-\kappa\Delta))\right)^{-\frac{2\kappa\theta}{\sigma^2}} \\ &\times \exp\left(\frac{iy \exp(-\kappa\Delta) r}{1 - \frac{\sigma^2}{2\kappa}iy(1 - \exp(-\kappa\Delta))}\right) \end{aligned} \quad (2)$$

2. *Die bedingte Dichte ist*

$$f_r(x|r_s = r) = \frac{1}{2b} \exp\left(-\frac{z + \delta}{2}\right) \left(\frac{z}{\delta}\right)^{\frac{\lambda}{4} - \frac{1}{2}} I_{\frac{\lambda}{2} - 1}(\sqrt{\delta z}) \quad (3)$$

Hierbei bezeichnet

$$b = \frac{\sigma^2}{4\kappa}(1 - \exp(-\kappa\Delta)), \quad z = \frac{x}{b}, \quad \lambda = \frac{4\kappa\theta}{\sigma^2}, \quad \delta = \frac{r}{b} \exp(-\kappa\Delta),$$

und  $I$  ist die modifizierte Bessel-Funktion erster Art.

*Beweis.* Um die bedingte Verteilung von  $r$  zu beschreiben, betrachten wir die charakteristische Funktion des Prozesses:

Für  $y \in \mathbb{R}$  ist die charakteristische Funktion  $\phi_X(y)$  einer Zufallsvariable  $X$  mit Verteilung  $F$  definiert durch

$$\phi_X(y) := \mathbb{E}[\exp(iyX)] = \int_{-\infty}^{\infty} \exp(iyx) dF(x)$$

Für den CIR Prozess betrachten wir nun wie in [10] den folgenden exponentiell-affinen Ansatz für die charakteristische Funktion, wobei  $u = iy$ :

$$\mathbb{E}[\exp(iyr_t)] = \mathbb{E}[\exp(ur_t)] = \exp(\phi(t, u) + \psi(t, u)r_0) \quad (4)$$

Wir suchen genauer gesagt Funktionen  $\phi$  und  $\psi$ , für die gilt, dass  $\phi(0, u) = 0$  und  $\psi(0, u) = u$  und dass  $M_t$ , definiert durch:

$$M_t = f(t, r_t) = \exp(\phi(T - t, u) + \psi(T - t, u)r_t) \quad (5)$$

ein Martingal ist. Denn damit folgt, dass

$$\mathbb{E}[\exp(iy r_T)] = \mathbb{E}[M_T] = M_0 = \exp(\phi(T, u) + \psi(T, u)r_0)$$

Mit diesen Überlegungen wenden wir nun auf (5) unter der Annahme, dass  $\psi$  und  $\phi$  ausreichend oft differenzierbar sind, die Ito-Formel an. Die dazu benötigten Ableitungen sind:

$$\begin{aligned} \frac{\partial}{\partial t} f(t, r_t) &= - \left( \frac{\partial}{\partial t} \phi(T - t, u) + r_t \frac{\partial}{\partial t} \psi(T - t, u) \right) f(t, r_t) \\ \frac{\partial}{\partial x} f(t, r_t) &= \psi(T - t, u) f(t, r_t) \\ \frac{\partial^2}{\partial x^2} f(t, r_t) &= \psi(T - t, u)^2 f(t, r_t) \end{aligned}$$

Ito Formel:

$$df(t, X_t) = \frac{\partial}{\partial t} f(t, X_t) dt + \frac{\partial}{\partial x} f(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, X_t) d[X_t]$$

Anwendung der Ito Formel auf (5) liefert:

$$\begin{aligned} \frac{df(t, r_t)}{f(t, r_t)} &= - \left( \frac{\partial}{\partial t} \phi(T - t, u) + r_t \frac{\partial}{\partial t} \psi(T - t, u) \right) dt + \psi(T - t, u) dr_t \\ &\quad + \frac{1}{2} \psi(T - t, u)^2 \sigma^2 r_t dt \\ &= - \left( \frac{\partial}{\partial t} \phi(T - t, u) + r_t \frac{\partial}{\partial t} \psi(T - t, u) \right) dt + \psi(T - t, u) \kappa(\theta - r_t) dt \\ &\quad + \psi(T - t, u) \sigma \sqrt{r_t} dW_t + \frac{1}{2} \psi(T - t, u)^2 \sigma^2 r_t dt \end{aligned}$$

Damit  $f(t, r_t)$  ein lokales Martingal ist, müssen die  $dt$  Terme einander aufheben, es muss also für alle  $r_t$  gelten, dass

$$\frac{\partial}{\partial t} \phi(T - t, u) + r_t \frac{\partial}{\partial t} \psi(T - t, u) = \psi(T - t, u) \kappa(\theta - r_t) + \frac{1}{2} \psi(T - t, u)^2 \sigma^2 r_t$$

Nun sammeln wir die Koeffizienten ein und erhalten die folgenden gewöhnlichen Differentialgleichungen für  $\phi$  und  $\psi$ , wobei  $s = T - t$ :

$$\begin{aligned}\frac{\partial}{\partial t}\phi(s, u) &= \theta\kappa\psi(s, u), & \phi(0, u) &= 0 \\ \frac{\partial}{\partial t}\psi(s, u) &= -\kappa\psi(s, u) + \frac{\sigma^2}{2}\psi(s, u)^2, & \psi(0, u) &= u\end{aligned}$$

Die Gleichung für  $\psi$  ist eine *Riccatische Differentialgleichung*. Man kennt daher eine Lösungsmethode dafür. Die Lösungen für die beiden Differentialgleichungen sehen folgendermaßen aus:

$$\begin{aligned}\phi(s, u) &= -\frac{2\kappa\theta}{\sigma^2} \log\left(1 - \frac{\sigma^2}{2\kappa}u(1 - \exp(-\kappa s))\right) \\ \psi(s, u) &= \frac{u \exp(-\kappa s)}{1 - \frac{\sigma^2}{2\kappa}u(1 - \exp(-\kappa s))}\end{aligned}$$

Somit ist die charakteristische Funktion gegeben durch:

$$\begin{aligned}\mathbb{E}[\exp(iyr_t)|r_0 = r] &= \left(1 - \frac{\sigma^2}{2\kappa}iy(1 - \exp(-\kappa t))\right)^{-\frac{2\kappa\theta}{\sigma^2}} \\ &\quad \times \exp\left(\frac{iy \exp(-\kappa t) r}{1 - \frac{\sigma^2}{2\kappa}iy(1 - \exp(-\kappa t))}\right)\end{aligned}\tag{6}$$

und die bedingte charakteristische Funktion für  $s < t$  und  $\Delta = t - s$  ist gegeben durch:

$$\begin{aligned}\mathbb{E}[\exp(iyr_t)|r_s = r] &= \left(1 - \frac{\sigma^2}{2\kappa}iy(1 - \exp(-\kappa\Delta))\right)^{-\frac{2\kappa\theta}{\sigma^2}} \\ &\quad \times \exp\left(\frac{iy \exp(-\kappa\Delta) r}{1 - \frac{\sigma^2}{2\kappa}iy(1 - \exp(-\kappa\Delta))}\right)\end{aligned}\tag{7}$$

Betrachtet man in (7) anstatt  $r_t$  die charakteristische Funktion von  $\frac{r_t}{b}$  mit  $b = \frac{\sigma^2}{4\kappa}(1 - \exp(-\kappa\Delta))$  und vergleicht man mit [9], erkennt man, dass dies die charakteristische Funktion einer nichtzentralen Chi-Quadrat Verteilung ist. Genauer ist  $\frac{1}{b}r_t|r_s$  nichtzentral Chi-Quadrat verteilt mit

$$\lambda = \frac{4\kappa\theta}{\sigma^2}, \delta = \frac{\exp(-\kappa\Delta) r}{b} \text{ wobei } b = \frac{\sigma^2}{4\kappa}(1 - \exp(-\kappa\Delta))\tag{8}$$

Hierbei sind  $\lambda$  die Freiheitsgrade und  $\delta$  ist der Nicht-Zentralitätsparameter. Da die charakteristische Funktion die Verteilung eindeutig charakterisiert, bedeutet das, dass  $\frac{1}{b}r_t|r_s$  nichtzentral Chi-Quadrat verteilt ist mit den Parametern  $\lambda$  und  $\delta$ .

Damit ist mit den Bezeichnungen wie in (8) und zudem  $z = \frac{x}{b}$  die bedingte Dichte wie behauptet (siehe [9])

$$f_r(x|r_s = r) = \frac{1}{2b} \exp\left(-\frac{z + \delta}{2}\right) \left(\frac{z}{\delta}\right)^{\frac{\lambda}{4} - \frac{1}{2}} I_{\frac{\lambda}{2} - 1}(\sqrt{\delta z})$$

□

Für spätere Berechnungen benötigen wir zudem die ersten vier bedingten Momente, insbesondere den bedingten Erwartungswert und die bedingte Varianz, und auch die bedingte Kovarianz:

**Satz 2.2.** *Sei*

$$b = \frac{\sigma^2}{4\kappa} (1 - \exp(-\kappa\Delta)), \quad \lambda = \frac{4\kappa\theta}{\sigma^2}, \quad \delta = \frac{\exp(-\kappa\Delta) r}{b}$$

*Damit gilt für  $s < t$ :*

$$\begin{aligned} \mathbb{E}[r_t|r_s = r] &= b[\lambda + \delta] \\ \mathbb{E}[r_t^2|r_s = r] &= b^2 [(\lambda + \delta)^2 + 2(\lambda + 2\delta)] \\ \mathbb{E}[r_t^3|r_s = r] &= b^3 [(\lambda + \delta)^3 + 6(\lambda + \delta)(\lambda + 2\delta) + 8(\lambda + 3\delta)] \\ \mathbb{E}[r_t^4|r_s = r] &= b^4 [(\lambda + \delta)^4 + 12(\lambda + \delta)^2(\lambda + 2\delta) + 4(11\lambda^2 + 44\lambda\delta + 36\delta^2) \\ &\quad + 48(\lambda + 4\delta)] \end{aligned} \tag{9}$$

*Insbesondere vereinfachen sich die Ausdrücke für den Erwartungswert und die bedingte Varianz zu*

$$\begin{aligned} \mathbb{E}[r_t|r_s = r] &= r \exp(-\kappa\Delta) + \theta(1 - \exp(-\kappa\Delta)) \\ \mathbb{V}[r_t|r_s = r] &= r \frac{\sigma^2}{\kappa} (\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) + \theta \frac{\sigma^2}{2\kappa} (1 - \exp(-\kappa\Delta))^2 \end{aligned} \tag{10}$$

*Die bedingte Kovarianz ist gegeben durch:*

$$\begin{aligned} \text{Cov}[r_t r_{t+h}|r_s = r] &= \exp(-\kappa h) \mathbb{E}[r_t^2|r_s = r] + \theta(1 - \exp(-\kappa h)) \mathbb{E}[r_t|r_s = r] \\ &\quad - \mathbb{E}[r_t|r_s = r] \mathbb{E}[r_{t+h}|r_s = r] \end{aligned} \tag{11}$$

*Beweis.* Die ersten vier Momente der nichtzentralen Chi-Quadrat-Verteilung können gewonnen werden durch Differenzieren der entsprechenden momenterzeugenden Funktion, welche folgendermaßen aussieht (siehe [9]):

$$M(t; \lambda, \delta) = \frac{1}{(1 - 2t)^{\frac{\lambda}{2}}} \exp\left(\frac{\delta t}{1 - 2t}\right)$$

Die Resultate für den bedingten Erwartungswert und die bedingte Varianz folgen sofort aus den Resultaten über die Momente durch Einsetzen von  $b$ ,  $\lambda$  und  $\delta$  und sind dessen vereinfachte Ausdrücke.

Für die Darstellung der bedingten Kovarianz stellen wir zunächst folgende Überlegungen an, wobei wir Eigenschaften der bedingten Erwartung verwenden:

$$\begin{aligned} \mathbb{E}[r_t r_{t+h} | r_s] &= \mathbb{E}[\mathbb{E}[r_t r_{t+h} | r_t] | r_s] = \mathbb{E}[r_t \mathbb{E}[r_{t+h} | r_t] | r_s] \\ &= \mathbb{E}[r_t (r_t \exp(-\kappa h) + \theta(1 - \exp(-\kappa h))) | r_s] \\ &= \exp(-\kappa h) \mathbb{E}[r_t^2 | r_s] + \theta(1 - \exp(-\kappa h)) \mathbb{E}[r_t | r_s] \end{aligned}$$

Damit folgt die behauptete Darstellung für die bedingte Kovarianz:

$$\begin{aligned} \text{Cov}[r_t r_{t+h} | r_s] &= \mathbb{E}[r_t r_{t+h} | r_s] - \mathbb{E}[r_t | r_s] \mathbb{E}[r_{t+h} | r_s] \\ &= \exp(-\kappa h) \mathbb{E}[r_t^2 | r_s] + \theta(1 - \exp(-\kappa h)) \mathbb{E}[r_t | r_s] \\ &\quad - \mathbb{E}[r_t | r_s] \mathbb{E}[r_{t+h} | r_s] \end{aligned}$$

□

## 2.2 Stationäre Verteilung

Bei Betrachtung der charakteristischen Funktion aus (6) für  $t \rightarrow \infty$ , sehen wir, dass  $t$  nur in den Termen  $\exp(-\kappa t)$  vorkommt und diese mit dem Grenzübergang Null werden, was zu folgender charakteristischen Funktion für die stationäre Verteilung führt:

$$\begin{aligned} &\lim_{t \rightarrow \infty} \left(1 - \frac{\sigma^2}{2\kappa} iy(1 - \exp(-\kappa t))\right)^{-\frac{2\kappa\theta}{\sigma^2}} \exp\left(\frac{iy \exp(-\kappa t) r}{1 - \frac{\sigma^2}{2\kappa} iy(1 - \exp(-\kappa t))}\right) \\ &= \left(1 - \frac{\sigma^2}{2\kappa} iy(1 - 0)\right)^{-\frac{2\kappa\theta}{\sigma^2}} \exp(0) = \left(1 - \frac{iy}{\frac{2\kappa}{\sigma^2}}\right)^{-\frac{2\kappa\theta}{\sigma^2}} \end{aligned}$$

Dies ist die charakteristische Funktion einer Gamma Verteilung mit Parametern

$$\nu = \frac{2\kappa\theta}{\sigma^2}, \quad \alpha = \frac{2\kappa}{\sigma^2} \quad (12)$$

Die Dichte sieht also folgendermaßen aus:

$$f_r(x) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} \exp(-\alpha x) \quad (13)$$

Für den stationären Erwartungswert und die stationäre Varianz ergeben sich somit mit  $X \sim \Gamma(\nu, \alpha)$

$$\mathbb{E}[X] = \theta \quad \mathbb{V}[X] = \frac{\sigma^2\theta}{2\kappa} \quad (14)$$

## 2.3 Simulation

Für die Schätzung der drei Parameter  $\kappa$ ,  $\theta$  und  $\sigma$  wurde zuerst ein Pfad eines Zinsverlaufes mit  $N$  Beobachtungen unter den wahren Parametern simuliert.

Das R-System stellt mit `rchisq` eine Funktion bereit, die Zufallsvariablen von einer nichtzentralen Chi-Quadrat Verteilung liefert. Damit generieren wir (ohne Diskretisierungsfehler, aus der exakten Übergangsverteilung (3)) einen Pfad  $x_1, \dots, x_N$  des CIR-Prozesses auf einem äquidistanten Gitter

$$x_i = r_{t_i}, \quad t_i = t_{i-1} + \Delta, \quad i = 1, \dots, N. \quad (15)$$

---

### Simulation CIR

---

ka=0.4

th=6

si=0.3

N=1000

**dt=1**

**b <- si ^ 2 / (4 \* ka) \* (1 - exp(-ka \* dt))**

```

x=vector(length=N)
x[1]=6

for(j in 2:N){
  x[j]=b*rchisq(1,4*ka*th/si^2,x[j-1]/b*exp(-ka*dt))
}

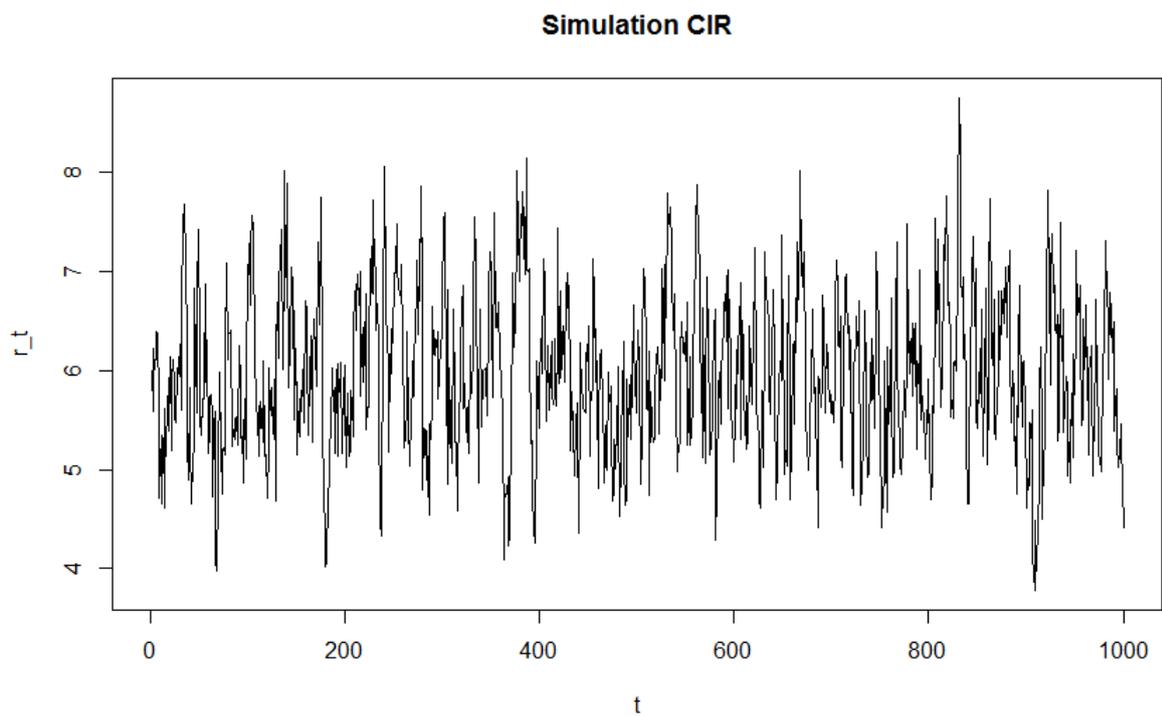
```

---

Führt man die Simulation mit dem Code von oben aus mit

$$\kappa = 0.4, \theta = 6, \sigma = 0.3, \Delta = 1, N = 1000$$

erhält man:



### 3 Parameterschätzung CIR Modell

Wir haben nun alle Information beisammen, um mit der eigentlichen Parameterschätzung zu beginnen. Dafür wurden die folgenden vier Methoden herangezogen:

- Maximum-Likelihood Methode
- Momentenmethode
- Exponentielle Momentenmethode
- Quasi-Maximum-Likelihood Methode

Hierbei ist im Folgenden immer der Parameterraum  $\Gamma = \{\gamma = (\kappa, \theta, \sigma) \in \mathbb{R}^3 : \kappa, \theta, \sigma > 0\}$  und der Parametervektor mit den wahren Parametern sei  $\gamma_0$ . Die  $N$  Beobachtungen, die unter  $\gamma_0$  simuliert wurden, werden mit  $r_1, \dots, r_N$  bezeichnet. Wir werden viele Schätzer iterativ aus entsprechenden Gleichungssystemen ermitteln. Für die Startwerte für diese Iterationen ziehen wir die folgenden einfachen unbedingten Momentenschätzer heran, die sich mithilfe der Formeln aus (10), (11) ergeben, wobei  $s < t$ :

$$\begin{aligned}\mathbb{E}[r_t] &= \theta \exp(-\kappa\Delta) + \theta(1 - \exp(-\kappa\Delta)) = \theta \\ \mathbb{V}[r_t] &= \theta \frac{\sigma^2}{\kappa} (\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) + \theta \frac{\sigma^2}{2\kappa} (1 - \exp(-\kappa\Delta))^2 = \frac{\theta\sigma^2}{2\kappa} \\ \text{Cov}[r_t r_{t+\Delta}] &= \exp(-\kappa\Delta) \left( \theta^2 + \frac{\theta\sigma^2}{2\kappa} \right) + \theta(1 - \exp(-\kappa\Delta))\theta - \theta^2 \\ &= \exp(-\kappa\Delta) \frac{\theta\sigma^2}{2\kappa} \Rightarrow \rho(1) = \frac{\text{Cov}[r_t r_{t+\Delta}]}{\sqrt{\mathbb{V}[r_t]} \sqrt{\mathbb{V}[r_{t+h}]} = \exp(-\kappa\Delta)\end{aligned}$$

Somit sind  $\kappa, \theta$  und  $\sigma$ :

$$\begin{aligned}\kappa &= -\frac{1}{\Delta} \log(\hat{\rho}(1)), \quad \text{wobei } \hat{\rho}(1) = \frac{\sum_{j=1}^{N-1} (r_{j+1} - \theta)(r_j - \theta)}{\sum_{j=1}^N (r_j - \theta)^2} \\ \theta &= \frac{1}{N} \sum_{j=1}^N r_j \\ \sigma &= \sqrt{\frac{2\kappa}{\theta} \frac{1}{N} \sum_{j=1}^N (r_j - \theta)^2}\end{aligned} \tag{16}$$

### 3.1 Maximum-Likelihood Methode

Es bietet sich an bei diesem Modell die Maximum-Likelihood Methode zu benutzen, da wir die Verteilung und insbesondere die bedingten Dichten exakt kennen und somit durch unsere Überlegungen von vorher sofort die zu maximierende bedingte log-likelihood Funktion anschreiben können:

$$l_N(\gamma) = \sum_{j=2}^N \log(g(r_j, r_{j-1}; \gamma)) \quad (17)$$

wobei  $g(r_j, r_{j-1}; \gamma) := \log(f_r(r_j | r_{j-1}; \gamma))$  ist.

Hierbei deutet das  $N$  die Abhängigkeit von der Anzahl der Beobachtungen und  $\gamma$  die Abhängigkeit vom Parametervektor.  $f_r$  bezeichnet die bedingte Dichte des CIR Prozesses, siehe (3).

Man beachte, dass die hier angegebene Funktion  $l_N(\gamma)$  nicht die klassische log-likelihood Funktion ist, da der erste Term, der nur vom Anfangswert  $r_1$  abhängt, weggelassen wurde. Diesen ersten Summanden zu vernachlässigen stellt statistisch gesehen bei großem  $N$  kein Problem dar, wie in [3] argumentiert wird.

Der dazugehörige R Code, der diese Methode implementiert, sieht folgendermaßen aus, wobei `ka1`, `th1` und `si1` im Code die Startwerte in (16) sind:

---

#### Maximum-Likelihood Methode CIR

---

```
logll <-function(ka, th, si){
  tmp <-0
  b <-si^2/(4*ka)*(1-exp(-ka*dt))
  for(j in 1:(N-1)){
    tmp <-tmp+log(dchisq(x[j+1]/b, 4*ka*th/si^2,
                        x[j]*exp(-ka*dt)/b))
  }
  return((-1)*tmp)
}
```

```
library(stats4)
```

```
optML <-mle(logl1 , start=list(ka=ka1 , th=th1 , si=si1 ) ,
           method = "L-BFGS-B" , lower=c(0.01 , 0.01 , 0.01) ,
           upper = c(Inf , Inf , Inf))
```

---

Wie man sieht, wurde die Maximierung in R mit der bereits in R implementierten Funktion `mle` vorgenommen, die für das Problem des Schätzens von Parametern mit der Maximum-Likelihood Methode vorgesehen ist. Als Optimierungsmethode wurde 'L-BFGS-B' gewählt, eine Modifikation der 'BFGS' quasi Newton-Methode, die obere und untere Schranken als zusätzliche Argumente zulässt. Dies ist in diesem Fall sinnvoll, da sämtliche Parameter größer als Null sein müssen.

### 3.2 Momentenmethode

Bei dieser einfacheren Methode betrachten wir die folgenden Martingal Schätzfunktionen, von der Konstruktion wie bei [8]:

$$\begin{aligned}
G_N^1(\gamma) &= \sum_{j=1}^N [r_j - f^1(r_{j-1}, \gamma)], & f^1(r, \gamma) &= \mathbb{E}_\gamma[r_1 | r_0 = r] \\
G_N^2(\gamma) &= \sum_{j=1}^N [r_j r_{j-1} - f^2(r_{j-1}, \gamma)], & f^2(r, \gamma) &= \mathbb{E}_\gamma[r_1 r_0 | r_0 = r] \\
G_N^3(\gamma) &= \sum_{j=1}^N [r_j^2 - f^3(r_{j-1}, \gamma)], & f^3(r, \gamma) &= \mathbb{E}_\gamma[r_1^2 | r_0 = r]
\end{aligned} \tag{18}$$

Hierbei ist  $\mathbb{E}_\gamma$  der Erwartungswert bezüglich des Parametervektors  $\gamma$ . Mithilfe der Formeln in (10) berechnen sich  $f^1, f^2, f^3$  als:

$$\begin{aligned}
f^1(r, \gamma) &= \mathbb{E}_\gamma[r_1 | r_0 = r] = r \exp(-\kappa\Delta) + \theta(1 - \exp(-\kappa\Delta)) \\
f^2(r, \gamma) &= \mathbb{E}_\gamma[r_1 r_0 | r_0 = r] = r^2 \exp(-\kappa\Delta) + r\theta(1 - \exp(-\kappa\Delta)) \\
f^3(r, \gamma) &= \mathbb{E}_\gamma[r_1^2 | r_0 = r] = \mathbb{V}_\gamma[r_1 | r_0 = r] + (\mathbb{E}_\gamma[r_1 | r_0 = r])^2 \\
&= r \frac{\sigma^2}{\kappa} (\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) + \theta \frac{\sigma^2}{2\kappa} (1 - \exp(-\kappa\Delta))^2 + \\
&\quad + (r \exp(-\kappa\Delta) + \theta(1 - \exp(-\kappa\Delta)))^2
\end{aligned} \tag{19}$$

Den Schätzer  $\hat{\gamma}_N$  des Parametervektors  $\gamma = (\kappa, \theta, \sigma)$  erhalten wir, indem wir die drei Gleichungen (18) gleich Null setzen, also  $G_N^j = 0, j = 1, 2, 3$ . Wir können diese drei Gleichungen mit drei Unbekannten, den Parametern, explizit lösen. Damit dies übersichtlich bleibt, führen wir die folgenden Bezeichnungen ein:

$$\begin{aligned}\beta &:= \exp(-\kappa\Delta) \\ \xi_{1,N} &:= \frac{1}{N} \sum_{i=1}^N r_i, & \xi_{2,N} &:= \frac{1}{N} \sum_{i=1}^N r_i r_{i-1} & \xi_{3,N} &:= \frac{1}{N} \sum_{i=1}^N r_i^2 \\ v_{1,N} &:= \frac{1}{N} \sum_{i=1}^N r_{i-1}, & v_{2,N} &:= \frac{1}{N} \sum_{i=1}^N r_{i-1}^2\end{aligned}$$

Betrachten wir damit nun die Gleichungen  $G_N^i, i = 1, 2, 3$ , erhalten wir:

$$\begin{aligned}G_N^1(\gamma) &= \sum_{j=1}^N [r_j - f^1(r_{j-1}, \gamma)] = \sum_{j=1}^N [r_j - r_{j-1}\beta - \theta(1 - \beta)] \\ &= \sum_{j=1}^N r_j - \beta \sum_{j=1}^N r_{j-1} - N\theta(1 - \beta) = N[\xi_{1,N} - \beta v_{1,N} - \theta(1 - \beta)] \stackrel{!}{=} 0\end{aligned}$$

$$\begin{aligned}G_N^2(\gamma) &= \sum_{j=1}^N [r_j r_{j-1} - f^2(r_{j-1}, \gamma)] = \sum_{j=1}^N [r_j r_{j-1} - r_{j-1}^2 \beta - r_{j-1} \theta(1 - \beta)] \\ &= \sum_{j=1}^N r_j r_{j-1} - \beta \sum_{j=1}^N r_{j-1}^2 - \theta(1 - \beta) \sum_{j=1}^N r_{j-1} \\ &= N[\xi_{2,N} - \beta v_{2,N} - \theta(1 - \beta) v_{1,N}] \stackrel{!}{=} 0\end{aligned}$$

$$\begin{aligned}
G_N^3(\gamma) &= \sum_{j=1}^N [r_j^2 - f^3(r_{j-1}, \gamma)] = \sum_{j=1}^N \left[ r_j^2 - \left( \frac{\sigma^2}{4\kappa} (1-\beta) \right)^2 \cdot \left[ \left( \frac{4\kappa\theta}{\sigma^2} + \frac{4\kappa\beta r_{j-1}}{\sigma^2(1-\beta)} \right)^2 \right. \right. \\
&\quad \left. \left. + 2 \left( \frac{4\kappa\theta}{\sigma^2} + 2 \frac{4\kappa\beta r_{j-1}}{\sigma^2(1-\beta)} \right) \right] \right] \\
&= \sum_{j=1}^N \left[ r_j^2 - (1-\beta)^2 \left( \theta^2 + 2\theta \frac{\beta r_{j-1}}{1-\beta} + \frac{\beta^2 r_{j-1}^2}{(1-\beta)^2} \right) \right. \\
&\quad \left. - \frac{\sigma^2}{2\kappa} (1-\beta)^2 \left( \theta + 2 \frac{\beta r_{j-1}}{1-\beta} \right) \right] \\
&= \sum_{j=1}^N r_j^2 - N\theta^2(1-\beta)^2 - 2\theta\beta(1-\beta) \sum_{j=1}^N r_{j-1} - \beta^2 \sum_{j=1}^N r_{j-1}^2 \\
&\quad - N \frac{\sigma^2}{2\kappa} (1-\beta)^2 \theta - \frac{\sigma^2}{\kappa} (1-\beta) \sum_{j=1}^N r_{j-1} \\
&= N \left[ \xi_{3,N} - \theta^2(1-\beta)^2 - 2\theta\beta(1-\beta)v_{1,N} - \beta^2 v_{2,N} - \frac{\sigma^2}{2\kappa} (1-\beta)^2 \theta \right. \\
&\quad \left. - \frac{\sigma^2}{\kappa} (1-\beta)v_{1,N} \right] \stackrel{!}{=} 0
\end{aligned}$$

Wir lösen nun diese drei Gleichungen für die Unbekannten  $\kappa, \theta$  und  $\sigma$  und erhalten nach einfacher Rechnung:

$$\begin{aligned}
\beta &= \frac{\xi_{2,N} - \xi_{1,N}v_{1,N}}{v_{2,N} - v_{1,N}^2} \Rightarrow \kappa = -\frac{1}{\Delta} \log \left( \frac{\xi_{2,N} - \xi_{1,N}v_{1,N}}{v_{2,N} - v_{1,N}^2} \right), \quad \theta = \frac{\xi_{1,N} - \beta v_{1,N}}{1-\beta}, \\
\sigma &= \sqrt{\frac{\xi_{3,N} - \theta^2(1-\beta)^2 - 2\theta\beta(1-\beta)v_{1,N} - \beta^2 v_{2,N}}{\frac{1}{2\kappa}(1-\beta)^2 \theta + \frac{1}{\kappa}(1-\beta)v_{1,N}}}
\end{aligned} \tag{20}$$

Die Implementierung in R sieht so aus:

---

Momentenmethode CIR

---

```
xi1 <- mean(x[2:N])
xi2 <- mean(x[2:N]*x[1:(N-1)])
xi3 <- mean(x[2:N]^2)
v1 <- mean(x[1:(N-1)])
v2 <- mean(x[1:(N-1)]^2)
be <- (xi2 - xi1*v1)/(v2-v1^2)
ka_est <- -log(be)/dt
th_est <- (xi1-be*v1)/(1-be)
si_est <- ((xi3-th^2*(1-be)^2-2*th*be*(1-be)*v1-be^2*v2)/
            (1/(2*ka)*(1-be)^2*th+1/ka*(1-be)*v1))^(1/2)
```

---

### 3.3 Exponentielle Momentenmethode

Eine weitere Möglichkeit besteht darin statt der normalen Momente die exponentiellen Momente heranzuziehen und gleichartige Schätzfunktionen zu betrachten, nämlich

$$G_N^k(\gamma) = \sum_{j=1}^N [\exp(iy_k r_j) - f(r_{j-1}, \gamma)], \quad k = 1, 2, 3 \quad (21)$$

$$f(r, \gamma) = \mathbb{E}_\gamma[\exp(iy_k r_1) | r_0 = r],$$

wobei man sofort sieht, dass die Funktion  $f$  auf der rechten Seite die uns explizit bekannte bedingte charakteristische Funktion ist, siehe (2).

Wesentlich für die drei Schätzfunktionen ist die Wahl der  $y_k$ . Während theoretisch eine beliebige Wahl von  $y_k \neq 0$  zum gewünschten Ergebnis führen sollte, zeigte sich durch praktisches Testen, dass die Qualität der Schätzung sehr wohl deutlich von der Wahl der  $y_k$  abhängt und in den meisten Fällen ergaben sich in zahlreichen Experimenten numerische Schwierigkeiten bzw. schlechte Schätzwerte.

Während man für eine Schätzung aus genau einem simulierten Pfad durchaus noch die  $y_k$  geschickt wählen kann, zeigte sich durch wiederholtes Ausführen,

dass selbst die  $y_k$ , die für einen bestimmten Pfad für die Schätzung zielführend waren, bei einer weiteren Simulation (zu denselben wahren Parametern und der anschließenden Schätzung danach) ähnlich schlechte Ergebnisse wie bei beliebiger Wahl der  $y_k$  liefern.

Aus diesen Gründen wurde bei dieser Schätzung ein etwas unkonventioneller Zugang gewählt:

Zuerst wird die Funktion  $G_N^k(y_k)$ , das heißt der obige Schätzer  $G_N^k$  als Funktion von  $y_k$ , betrachtet, im Code genannt 'testy'. Die Parameter werden in dieser Funktion durch die Vorschätzer ersetzt, nämlich eben die in (16).

Wir wissen, dass theoretisch  $G_N^k(\gamma) = 0$  erfüllen muss, sofern man die wahren Parameter einsetzt. Im zweiten Schritt betrachten wir daher die Nullstellen der Funktion  $G_N^k(y_k)$ , um gute Werte für  $y_k$  zu finden.

Wichtig ist, sich zu erinnern, dass wir hierfür nicht die wahren Parameterwerte verwendet haben, sondern deren Vorschätzer, die direkt aus den simulierten Daten gewonnen werden. Würden wir die wahren Werte verwenden, hätten wir nämlich für unseren endgültigen Schätzer Information verwendet, die wir zu diesem Zeitpunkt nicht haben sollten - die Parameter zu schätzen unter Verwendung der wahren Parameter ergibt keinen Sinn.

Die Funktion  $G_N^k(y_k)$  hat bei beliebiger Simulation als Grundlage immer eine ähnliche Gestalt, ist wie erwartet symmetrisch mit trivialer Nullstelle bei 0. Es ist also einfach möglich drei nicht-triviale Werte für  $y_k$  unter diesen zu finden, selbstverständlich mit Vermeidung von symmetrischen Nullstellen.

Mit dieser automatisierten Wahl für  $y_1, y_2, y_3$  zeigt sich nun auch bei wiederholtem Testen mit vielen Pfaden, dass die Schätzung der Parameter durch diese angepasste Methode sinnvolle und vergleichbare Ergebnisse liefert.

Der folgende R-Code implementiert die beschriebene Methode:

---

Exponentielle Momentenmethode CIR

---

```
fexp <- function(r, ka, th, si, y){
  b <- si^2/(4*ka)*(1-exp(-ka*dt))
  return((1-1i*y*2*b)^(-2*ka*th/si^2)*
    exp(1i*y*exp(-ka*dt)*r/(1-1i*y*2*b)))
```

```

}

testy <- function(y1)
  Re(mean(exp(1i*y1*x[2:N]) -
           fexp(x[1:(N-1)],ka1,th1,si1,y1)))
testy1 <- Vectorize(testy)
yroot<-uniroot.all(testy1, c(-2,2))
y1 <- yroot[2]
y2 <- yroot[3]
y3 <- yroot[5]

library(rootSolve)

modelexp <- function(y)
  c(F1 = Re(mean(exp(1i*y1*x[2:N]) -
                 fexp(x[1:(N-1)],y[1],y[2],y[3],y1))),
    F2 = Re(mean(exp(1i*y2*x[2:N]) -
                 fexp(x[1:(N-1)],y[1],y[2],y[3],y2))),
    F3 = Re(mean(exp(1i*y3*x[2:N]) -
                 fexp(x[1:(N-1)],y[1],y[2],y[3],y3))) )

optExpMo <- multiroot(f=modelexp, start=c(ka1,th1,si1))

```

Das eigentliche Lösen dieses Gleichungssystems wurde hier mit der Funktion `multiroot` aus dem Paket `rootSolve` vorgenommen. Diese Funktion ist speziell für das Finden von Nullstellen bei mehrdimensionalen Funktionen vorgesehen.

### 3.4 Quasi-Maximum-Likelihood Methode

Während wir bei der Maximum-Likelihood Methode für die log-likelihood Funktion die wahren Übergangsdichten genommen haben, nehmen wir hier stattdessen normalverteilte Übergangsdichten an und schätzen Erwartungswert und Varianz dafür aus unseren simulierten Daten. Die log-likelihood

Funktion sieht also folgendermaßen aus:

$$l_N(\gamma) = \log(\eta(\bar{\mu}_1, \bar{\sigma}_1^2; \gamma)) + \sum_{j=2}^N \log(\eta(r_j, \mu(r_{j-1}), \sigma^2(r_{j-1}); \gamma)) \quad (22)$$

wobei  $\eta$  die Dichte der Normalverteilung bezeichnet.

Hierbei wurde für den ersten Summanden Erwartungswert und Varianz geschätzt aus der stationären Verteilung, nämlich einer Gamma Verteilung mit Parametern  $\nu = \frac{2\kappa\theta}{\sigma^2}, \omega = \frac{2\kappa}{\sigma^2}$ , siehe (12), und für alle weiteren Summanden erfolgte die Schätzung mit dem bedingten Erwartungswert und der bedingten Varianz aus (10). Das bedeutet also:

$$\begin{aligned} \bar{\mu}_1 &= \frac{\nu}{\omega} = \frac{\frac{2\kappa\theta}{\sigma^2}}{\frac{2\kappa}{\sigma^2}} = \theta \\ \bar{\sigma}_1^2 &= \frac{\nu}{\omega^2} = \frac{\frac{2\kappa\theta}{\sigma^2}}{\frac{4\kappa^2}{\sigma^4}} = \frac{\sigma^2\theta}{2\kappa} \\ \mu(r_{j-1}) &= \mathbb{E}_\gamma[r_j | r_{j-1} = r] \\ \sigma^2(r_{j-1}) &= \mathbb{V}_\gamma[r_j | r_{j-1} = r] \end{aligned} \quad (23)$$

Diese Methode wird durch den folgenden Code implementiert:

---

Quasi-Maximum-Likelihood Methode CIR

```

a1 <- function(r, ka, th, si){
  r*exp(-ka*dt) + th*(1-exp(-ka*dt))
}

b1 <- function(r, ka, th, si){
  r*(si^2/ka)*(exp(-ka*dt) - exp(-2*ka*dt)) +
  th*(si^2/(2*ka))*(1-exp(-ka*dt))^2
}

g <- function(ka, th, si){
  mul <- th

```

```

si1 <- si^2*th/(2*ka)
gstart <- log(dnorm(x[1], mu1, si1))
for (j in 2:N){
  ghelp <- log(dnorm(x[j], a1(x[j-1], ka, th, si),
                    b1(x[j-1], ka, th, si)))
  gstart <- gstart+ghelp
}
return((-1)*gstart)
}

```

```
library(stats4)
```

```
optQML<-mle(g, start=list(ka=ka1, th=th1, si=si1))
```

---

Die Funktionen  $a1$  und  $b1$  implementieren die obigen bedingten Erwartungswerte und bedingten Varianzen  $\mu(r_{j-1})$  bzw.  $\sigma^2(r_{j-1})$ . Wie bei der Maximum-Likelihood Methode wird die Schätzung hier ebenfalls mit der Funktion `mle` vorgenommen.

## 4 Varianzschätzung CIR Modell

### 4.1 Varianzschätzung Maximum-Likelihood Methode

Sei  $g(r_j, r_{j-1}; \gamma) := \log(f_r(r_j|r_{j-1}; \gamma))$ . Damit ergibt sich als einfacher Schätzer für die Varianzen (vgl. [3]):

$$\left( \sum_{j=2}^N \frac{\partial}{\partial \gamma} g(r_j, r_{j-1}; \gamma) \frac{\partial}{\partial \gamma} g(r_j, r_{j-1}; \gamma)^T \right)^{-1}$$

Für die Berechnung erinnern wir uns, dass die Übergangsdichte folgendermaßen aussieht (siehe (3)), wobei  $h$  die Dichte einer nichtzentralen Chi-Quadrat Verteilung bezeichnet:

$$f_r(r_j|r_{j-1}; \gamma) = \frac{1}{b} h \left( \frac{r_j}{b}, \frac{4\kappa\theta}{\sigma^2}, \frac{\exp(-\kappa\Delta) r_{j-1}}{b} \right) \text{ mit } b := \frac{\sigma^2}{4\kappa} (1 - \exp(-\kappa\Delta))$$

Unter Beachtung der Darstellung der nichtzentralen Chi-Quadrat Dichte ergibt sich für  $g$  die folgende Darstellung:

$$\begin{aligned} g(r_j, r_{j-1}; \gamma) &:= \log(f_r(r_j|r_{j-1}; \gamma)) = -\log b + \log h \left( \frac{r_j}{b}, \frac{4\kappa\theta}{\sigma^2}, \frac{\exp(-\kappa\Delta) r_{j-1}}{b} \right) \\ &= -\log b - \log 2 - \frac{1}{2b} (r_j + \exp(-\kappa\Delta) r_{j-1}) + \left( \frac{1}{4} \frac{4\kappa\theta}{\sigma^2} - \frac{1}{2} \right) \\ &\quad \times \log \left( \frac{r_j}{\exp(-\kappa\Delta) r_{j-1}} \right) + \log I_{\frac{4\kappa\theta}{2\sigma^2}-1} \left( \sqrt{\frac{\exp(-\kappa\Delta) r_j r_{j-1}}{b^2}} \right) \\ &= -2 \log(\sigma) + \log(4) + \log(\kappa) - \log(1 - \exp(-\kappa\Delta)) - \log(2) \\ &\quad - \frac{2\kappa}{\sigma^2(1 - \exp(-\kappa\Delta))} (r_j + \exp(-\kappa\Delta)r_{j-1}) + \left( \frac{\kappa\theta}{\sigma^2} - \frac{1}{2} \right) \\ &\quad \times (\log(r_j) - \log(r_{j-1}) + \kappa\Delta) \\ &\quad + \log I_{\frac{4\kappa\theta}{2\sigma^2}-1} \left( \sqrt{\frac{16\kappa^2 \exp(-\kappa\Delta) r_j r_{j-1}}{\sigma^4(1 - \exp(-\kappa\Delta))^2}} \right) \end{aligned} \tag{24}$$

Die Ableitungen von  $g$  nach den drei Parametern sehen damit folgendermaßen aus:

$$\begin{aligned}
\frac{\partial}{\partial \kappa} g(r_j, r_{j-1}; \gamma) &= \frac{1}{\kappa} - \frac{\Delta \exp(-\kappa \Delta)}{1 - \exp(-\kappa \Delta)} - \frac{2}{\sigma^2} \frac{(r_j + \exp(-\kappa \Delta) r_{j-1})(1 - \Delta \kappa)}{(1 - \exp(-\kappa \Delta))} \\
&+ \frac{2}{\sigma^2} \frac{\Delta \kappa \exp(-\kappa \Delta) (r_j + \exp(-\kappa \Delta) r_{j-1})}{(1 - \exp(-\kappa \Delta))^2} \\
&+ \frac{\theta}{\sigma^2} (\log(r_j) - \log(r_{j-1}) + \kappa \Delta) \\
&+ \Delta \left( \frac{\kappa \theta}{\sigma^2} - \frac{1}{2} \right) + \frac{\partial}{\partial \kappa} \log I_{\frac{4\kappa\theta}{2\sigma^2}-1} \left( \sqrt{\frac{16\kappa^2 \exp(-\kappa \Delta) r_j r_{j-1}}{\sigma^4 (1 - \exp(-\kappa \Delta))^2}} \right) \\
\frac{\partial}{\partial \theta} g(r_j, r_{j-1}; \gamma) &= \frac{\kappa}{\sigma^2} (\log(r_j) - \log(r_{j-1}) + \kappa \Delta) \\
&+ \frac{\partial}{\partial \theta} \log I_{\frac{4\kappa\theta}{2\sigma^2}-1} \left( \sqrt{\frac{16\kappa^2 \exp(-\kappa \Delta) r_j r_{j-1}}{\sigma^4 (1 - \exp(-\kappa \Delta))^2}} \right) \\
\frac{\partial}{\partial \theta} g(r_j, r_{j-1}; \gamma) &= -\frac{2}{\sigma} + \frac{4\kappa(r_j + \exp(-\kappa \Delta) r_{j-1})}{(1 - \exp(-\kappa \Delta)) \sigma^3} - \frac{2\kappa\theta}{\sigma^3} (\log(r_j) - \log(r_{j-1})) \\
&+ \kappa \Delta + \frac{\partial}{\partial \sigma} \log I_{\frac{4\kappa\theta}{2\sigma^2}-1} \left( \sqrt{\frac{16\kappa^2 \exp(-\kappa \Delta) r_j r_{j-1}}{\sigma^4 (1 - \exp(-\kappa \Delta))^2}} \right)
\end{aligned} \tag{25}$$

Schließlich benötigen wir noch die Ableitung der logarithmierten Besselfunktion. Betrachten wir allgemein eine Besselfunktion mit Index  $\nu$  und Argument  $x$ , ergibt sich für die Ableitung nach  $x$ :

$$\frac{\partial}{\partial x} \log I_\nu(x) = \frac{I_{\nu-1}(x)}{I_\nu(x)} - \frac{\nu}{x}$$

Für die Ableitung nach  $\nu$  verwenden wir die folgende Darstellung der Besselfunktion aus [1]:

$$I_\nu(x) = \left(\frac{x}{2}\right)^\nu \frac{1}{\sqrt{\pi} \Gamma(\nu + \frac{1}{2})} \int_0^\pi \exp(x \cos(h)) (\sin(h))^{2\nu} dh$$

woraus sich für uns die gewünschte Ableitung dann ergibt als:

$$\frac{\partial}{\partial \nu} \log I_\nu(x) = \log\left(\frac{x}{2}\right) - \psi\left(\nu + \frac{1}{2}\right) + \frac{\int_0^\pi \exp(x \cos(h)) (\sin(h))^{2\nu} \log(\sin(h)) dh}{\int_0^\pi \exp(x \cos(h)) (\sin(h))^{2\nu} dh}$$

wobei  $\psi(s)$  die Digamma Funktion bezeichnet, also  $\psi(s) = \frac{d}{ds} \log \Gamma(s) = \frac{\Gamma'(s)}{\Gamma(s)}$ .

## 4.2 Varianzschätzung Momentenmethode

Eine Vorgehensweise, um die Varianzen der Parameter theoretisch zu schätzen, wird beschrieben in [8] und wir gehen analog vor:

Zunächst benötigen wir folgenden Martingalkonvergenzsatz (entspricht Lemma 3.8 in [8]):

**Satz 4.1.** *Angenommen  $(X_{n,k})$  ist ein Martingaldifferenzvektor, sodass für alle  $\epsilon > 0$  gilt, dass*

$$\sum_{k=1}^n \mathbb{E}[\|X_{n,k}\|^2 \mathbb{1}_{\{\|X_{n,k}\| > \epsilon\}} | \mathcal{F}_{k-1}] \xrightarrow{P} 0$$

und

$$\sum_{k=1}^n [X_{n,k} X_{n,k}^T | \mathcal{F}_{k-1}] \xrightarrow{P} Y$$

mit  $n \rightarrow \infty$ . Dann gilt:

$$\sum_{k=1}^n X_{n,k} \xrightarrow{D} N(0, Y)$$

**Satz 4.2.** *In unserem speziellen Fall gilt damit (entspricht Proposition 3.9 in [8]):*

$$\frac{1}{\sqrt{N}} G_N(\gamma_0) \xrightarrow{D} N(0, Y)$$

mit  $N \rightarrow \infty$ , wobei die Matrix  $Y$  folgendermaßen aussieht:

$$Y_{ij} = \mathbb{E}(\text{Cov}[\Xi_k^i, \Xi_k^j | r_0])$$

Hierbei ist  $\Xi_k = (r_k, r_k r_{k-1}, r_k^2)^T$ .

Explizites Ausrechnen der Matrix  $Y$  ergibt unter Verwendung von (9):

### Matrix $Y$

$$\begin{aligned} Y_{11} &= \mathbb{E}(\text{Cov}[r_1, r_1|r_0]) = \mathbb{V}[r_1|r_0] \\ Y_{12} = Y_{21} &= \mathbb{E}(\text{Cov}[r_1, r_1 r_0|r_0]) = \mathbb{E}[r_0]\mathbb{V}[r_1|r_0] = \mathbb{E}[r_0]Y_{11} \\ Y_{13} = Y_{31} &= \mathbb{E}(\text{Cov}[r_1, r_1^2|r_0]) = \mathbb{E}[r_1^3|r_0] - \mathbb{E}[r_1|r_0]\mathbb{E}[r_1^2|r_0] \\ Y_{22} &= \mathbb{E}(\text{Cov}[r_1 r_0, r_1 r_0|r_0]) = \mathbb{E}[r_0^2]\mathbb{V}[r_1|r_0] = \mathbb{E}[r_0^2]Y_{11} \\ Y_{23} = Y_{32} &= \mathbb{E}(\text{Cov}[r_1 r_0, r_1^2|r_0]) = \mathbb{E}[r_0]\mathbb{E}(\text{Cov}[r_1, r_1^2|r_0]) = \mathbb{E}[r_0]Y_{13} \\ Y_{33} &= \mathbb{E}(\text{Cov}[r_1^2, r_1^2|r_0]) = \mathbb{E}[r_1^4|r_0] - (\mathbb{E}[r_1^2|r_0])^2 \end{aligned}$$

Im zweiten Schritt verwenden wir (entspricht Lemma 3.11 in [8]):

**Satz 4.3.** *Es gilt*

$$\sqrt{N}[\xi_N - \xi] \xrightarrow{D} N(0, \Sigma)$$

wobei

$$\Sigma = P^{-1}Y(P^{-1})^T$$

und

$$P_{ij} = \delta_{ij} - \phi_1^i \delta_{1j} - \phi_2^i \delta_{3j}$$

Hierbei bezeichnet  $\delta_{ij}$  das Kroneker Delta und die  $\phi_1^i$  erhalten wir aus der folgenden Darstellung der Funktionen  $f^i(r, \gamma)$ ,  $i = 1 \dots 3$ :

$$f^i(r, \gamma) = \phi_2^i r^2 + \phi_1^i r + \phi_0^i$$

Explizites Ausrechnen der Matrix  $P$  ergibt unter Verwendung von (19):

### Matrix $P$

$$\begin{aligned} P_{11} &= 1 - \exp(-\kappa\Delta) \\ P_{12} &= 0 \\ P_{13} &= 0 \\ P_{21} &= -\theta - \exp(-\kappa\Delta) \\ P_{22} &= 1 \\ P_{23} &= -\exp(-\kappa\Delta) \\ P_{31} &= -\frac{\sigma^2}{\kappa}(\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) + 2\exp(-\kappa\Delta)\theta(1 - \exp(-\kappa\Delta)) \\ P_{32} &= 1 - \exp(-\kappa\Delta) \\ P_{33} &= 1 - \exp(-2\kappa\Delta) \end{aligned}$$

Im letzten Schritt erhalten wir schließlich die Verteilung des Schätzers (entspricht Theorem 3.12 in [8]):

**Satz 4.4.** *Der Schätzer  $\hat{\gamma}_N = (\kappa_N, \theta_N, \sigma_N)$  ist asymptotisch normal, nämlich mit*

$$\sqrt{N}[\hat{\gamma}_N - \gamma_0] \xrightarrow{D} N(0, T)$$

mit  $N \rightarrow \infty$ , wobei

$$T = D\Sigma D^T$$

Hierbei sind die Einträge der Matrix  $D$  gegeben durch (26).

$$D_{ij} = \frac{\partial h_i}{\partial x_j}(\xi_N) \quad i = 1 \dots 3 \quad (26)$$

Die Funktionen  $h_i$  sind Schätzfunktionen der drei Parameter. Während wir beim Schätzen in (20) Funktionen von  $\xi_N = (\xi_{1,N}, \xi_{2,N}, \xi_{3,N})$  und  $v_N = (v_{1,N}, v_{2,N})$  betrachtet haben, was uns exakte Schätzer liefert, kann man durch Taylor Entwicklung argumentieren, dass  $\hat{\gamma}_N = h(\xi_N) + S_N$ , wobei wir in unseren neuen Schätzern  $h_i$  schlichtweg  $v_{1,N}$  und  $v_{2,N}$  durch  $\xi_{1,N}$  bzw.  $\xi_{3,N}$  ersetzen. Der Term  $S_N$  kann vernachlässigt werden.

Wieder führen wir die explizite Berechnung durch mithilfe von (20):

### Matrix D

Betrachten wir also die Funktionen  $h_1, h_2$  und  $h_3$ , dann sehen diese folgendermaßen aus:

$$\hat{\kappa} = h_1(\xi_N) = -\frac{1}{\Delta} \log \left( \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2} \right),$$

$$\hat{\theta} = h_2(\xi_N) = \frac{\xi_{1,N} - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2} \xi_{1,N}}{1 - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}} = \xi_{1,N},$$

$$\begin{aligned}\hat{\sigma} &= h_3(\xi) = \sqrt{\frac{\xi_{3,N} - \theta^2(1-\beta)^2 - 2\theta\beta(1-\beta)\xi_{1,N} - \beta^2\xi_{3,N}}{\frac{1}{2\kappa}(1-\beta)^2\theta + \frac{1}{\kappa}(1-\beta)\xi_{1,N}}} \\ &= \sqrt{\frac{\xi_{3,N} - \xi_{1,N}^2\left(1 - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\right)^2 - 2\xi_{1,N}\frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\left(1 - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\right)\xi_{1,N} - \left(\frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\right)^2 \xi_{3,N}}{\sqrt{-\frac{\Delta}{\log\left(\frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\right)}\left[\frac{1}{2}\left(1 - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\right)^2\xi_{1,N} + \left(1 - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}\right)\xi_{1,N}\right]}}\end{aligned}$$

Damit ergeben sich die gewünschten Ableitungen und somit die Einträge der Matrix  $D$  als:

$$\begin{aligned}D_{11} &= \frac{\partial}{\partial \xi_{1,N}} h_1(\xi_N) = \frac{1}{\Delta} \frac{2\xi_{1,N}}{\xi_{2,N} - \xi_{1,N}^2} - \frac{1}{\Delta} \frac{2\xi_{1,N}}{\xi_{3,N} - \xi_{1,N}^2} \\ D_{12} &= \frac{\partial}{\partial \xi_{2,N}} h_1(\xi_N) = -\frac{1}{\Delta} \frac{\xi_{1,N}}{\xi_{2,N} - \xi_{1,N}^2} \\ D_{13} &= \frac{\partial}{\partial \xi_{3,N}} h_1(\xi_N) = \frac{1}{\Delta} \frac{\xi_{1,N}}{\xi_{3,N} - \xi_{1,N}^2} \\ D_{21} &= \frac{\partial}{\partial \xi_{1,N}} h_2(\xi_N) = 1 \\ D_{22} &= \frac{\partial}{\partial \xi_{2,N}} h_2(\xi_N) = 0 \\ D_{23} &= \frac{\partial}{\partial \xi_{3,N}} h_2(\xi_N) = 0\end{aligned}$$

Während die Berechnung der Einträge von  $D_{31}$ ,  $D_{32}$  und  $D_{33}$ , was den Ableitungen der Funktion  $h_3$  entspricht, zwar theoretisch nicht schwer sind, ergeben sich jedoch sehr lange Ausdrücke. Deshalb werden diese Ausdrücke hier nicht aufgelistet.

Die Berechnung wurde zudem nicht händisch, sondern mit einem externen Programm durchgeführt und schließlich nach R zur weiteren Rechnung exportiert.

### 4.3 Bemerkung zur Varianzschätzung Exponentielle Momentenmethode

Aufgrund der Änderung dieser Methode, damit diese funktioniert, ist die bekannte Theorie hierbei nicht anwendbar.

## 4.4 Varianzschätzung Quasi-Maximum-Likelihood Methode

Im Falle der QML Methode ist analog zur ML Methode ein einfacher Schätzer für die Varianzen, vgl. [7]:

$$\left( \frac{\partial}{\partial \gamma} \log(\eta(\bar{\mu}_1, \bar{\sigma}_1^2; \gamma)) \frac{\partial}{\partial \gamma} \log(\eta(\bar{\mu}_1, \bar{\sigma}_1^2; \gamma))^T + \sum_{j=2}^N \frac{\partial}{\partial \gamma} g(r_j, r_{j-1}; \gamma) \frac{\partial}{\partial \gamma} g(r_j, r_{j-1}; \gamma)^T \right)^{-1}$$

wobei  $g(r_j, r_{j-1}; \gamma) := \log(\eta(r_j, \mu(r_{j-1}), \sigma^2(r_{j-1}); \gamma))$  für  $j > 1$  ist.  $\eta$  bezeichnet hier wie bereits bei der Parameterschätzung die Dichte der Normalverteilung.

Zum Bestimmen der Ableitungen betrachten wir zunächst eine allgemeine Normalverteilung  $\eta(\mu(y), \sigma^2(y))$  und leiten diese nach  $y$  ab:

$$\begin{aligned} \frac{\partial}{\partial y} \log(\eta(\mu(y), \sigma^2(y))) &= \frac{\partial}{\partial y} \log \left( \frac{1}{\sqrt{2\pi\sigma^2(y)}} \exp \left( -\frac{(x - \mu(y))^2}{2\sigma^2(y)} \right) \right) \\ &= \frac{\partial}{\partial y} \left[ \log \left( \frac{1}{\sqrt{2\pi\sigma^2(y)}} \right) + \left( -\frac{(x - \mu(y))^2}{2\sigma^2(y)} \right) \right] \\ &= \frac{\partial}{\partial y} \left[ -\log \left( \sqrt{2\pi\sigma^2(y)} \right) - \frac{(x - \mu(y))^2}{2\sigma^2(y)} \right] \\ &= \frac{\frac{\partial}{\partial y}[\sigma^2(y)]}{2\sigma^2(y)} + \frac{2(x - \mu(y)) \frac{\partial}{\partial y}[\mu(y)]\sigma^2(y) + (x - \mu(y))^2 \frac{\partial}{\partial y}[\sigma^2(y)]}{2(\sigma^2(y))^2} \end{aligned} \quad (27)$$

Nun brauchen wir nur noch die jeweiligen Ableitungen nach den Parametern der Funktionen  $\mu$  und  $\sigma^2$  zu bestimmen. Man beachte, dass dadurch, dass der erste Summand von der stationären Verteilung abhängt, man diesen auch in Bezug auf die Ableitungen separat betrachten muss, vergleiche (23).

$$j = 1 : \quad \begin{array}{llll} \mu_1 = \theta & \frac{\partial}{\partial \kappa} \mu_1 = 0 & \frac{\partial}{\partial \theta} \mu_1 = 1 & \frac{\partial}{\partial \sigma} \mu_1 = 0 \\ \sigma_1^2 = \frac{\sigma^2 \theta}{2\kappa} & \frac{\partial}{\partial \kappa} \sigma_1^2 = -\frac{\sigma^2 \theta}{2\kappa^2} & \frac{\partial}{\partial \theta} \sigma_1^2 = \frac{\sigma^2}{2\kappa} & \frac{\partial}{\partial \sigma} \sigma_1^2 = \frac{\sigma \theta}{\kappa} \end{array}$$

$$\begin{aligned}
j > 1 : \quad & \mu(r_j) = r \exp(-\kappa\Delta) + \theta(1 - \exp(-\kappa\Delta)) \\
& \frac{\partial}{\partial \kappa} \mu(r_j) = -r_j \exp(-\kappa\Delta)\Delta + \theta \exp(-\kappa\Delta)\Delta \\
& \frac{\partial}{\partial \theta} \mu(r_j) = 1 - \exp(-\kappa\Delta) \\
& \frac{\partial}{\partial \sigma} \mu(r_j) = 0 \\
& \sigma(r_j) = r \frac{\sigma^2}{\kappa} (\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) + \theta \frac{\sigma^2}{2\kappa} (1 - \exp(-\kappa\Delta))^2 \\
& \frac{\partial}{\partial \kappa} \sigma(r_j) = -\frac{r_j \sigma^2}{\kappa^2} (\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) - \frac{\theta \sigma^2}{2\kappa^2} (1 - \exp(-\kappa\Delta))^2 \\
& \quad + \frac{r_j \sigma^2}{\kappa} (2\Delta \exp(-2\kappa\Delta) - \Delta \exp(-\kappa\Delta)) \\
& \quad + \frac{\theta \sigma^2}{\kappa} (1 - \exp(-\kappa\Delta)) \exp(-\kappa\Delta) \Delta \\
& \frac{\partial}{\partial \theta} \sigma(r_j) = \frac{\sigma^2}{2\kappa} (1 - \exp(-\kappa\Delta))^2 \\
& \frac{\partial}{\partial \sigma} \sigma(r_j) = 2\sigma \frac{r_j}{\kappa} (\exp(-\kappa\Delta) - \exp(-2\kappa\Delta)) + 2\sigma \frac{\theta}{2\kappa} (1 - \exp(-\kappa\Delta))
\end{aligned}$$

## 5 Vergleich und numerische Illustrationen CIR Modell

Zum Vergleich der vier verschiedenen Methoden kann man einerseits für konkrete Parameterwerte dessen theoretische Varianzschätzer betrachten, andererseits kann man sich auch durch wiederholtes Schätzen der Parameter aus mehrfachen Simulationen die Schätzer für die jeweiligen Parameter berechnen und diese Daten dann zum Beispiel als Histogramm darstellen.

Im Folgenden wird beides für die jeweiligen Methoden präsentiert. Hierbei ist zu beachten, dass es bei der exponentiellen Momentenmethode keinen theoretischen Schätzer gibt, wie bereits im vorherigen Kapitel angemerkt.

Wie in der anfänglichen Illustration wählen wir folgende Parameter, wobei wir jetzt jedoch mehr Zeitschritte betrachten:

$$\kappa = 0.4, \theta = 6, \sigma = 0.3, \Delta = 1, N = 10000$$

Die Anzahl der Wiederholungen bei der Simulation und anschließendem Schätzen für die Histogramme ist  $M = 100$ .

### Maximum-Likelihood Methode

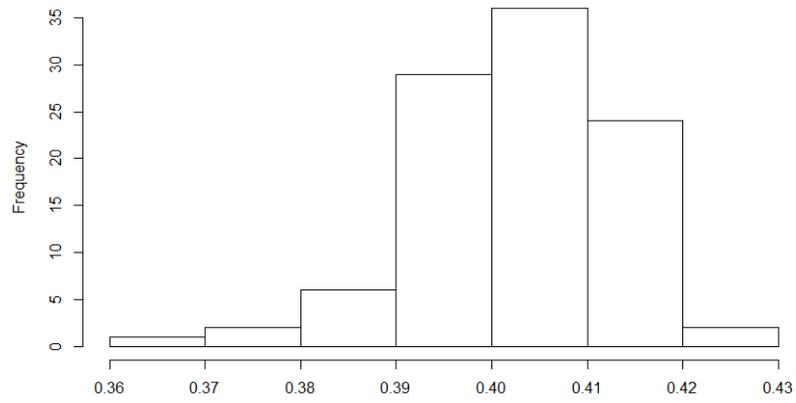
Bei dieser Methode merkt man, dass die Einträge der Varianz-Kovarianz Matrix sehr klein sind. Auch bei Betrachtung der Histogramme erkennt man eine sehr gute Schätzung der Parameter und nur geringe Abweichungen vom wahren Wert. Dafür ist die Berechnung mit dieser Methode auch sehr rechenaufwändig, was sich besonders bei der hundertmaligen Wiederholung für die Histogramme bemerkbar gemacht hat.

Der theoretische Varianzschätzer, berechnet mit der Methode aus Kapitel 4.1, ist durch folgende Matrix gegeben:

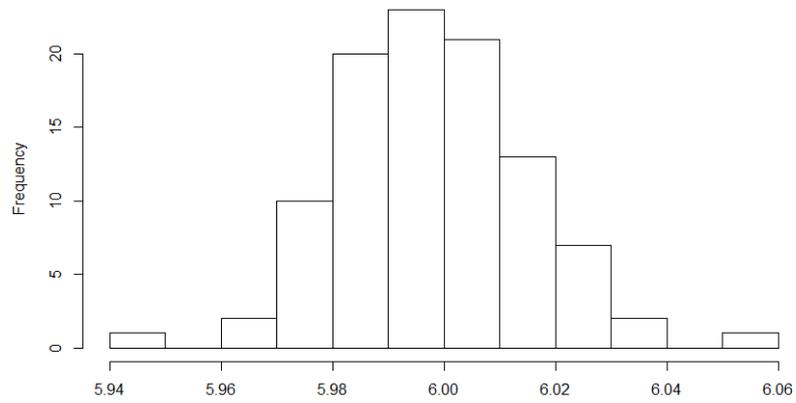
$$\begin{bmatrix} 0.075588098 & -0.54592318 & 0.003130164 \\ 0.545923180 & 5.09794452 & -0.016447047 \\ 0.003130164 & -0.01644705 & 0.000162763 \end{bmatrix}$$

Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:

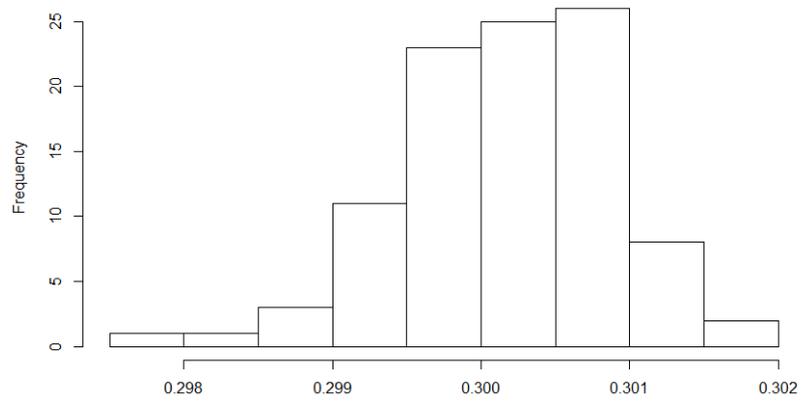
**CIR Maximum Likelihood Methode - kappa**



**CIR Maximum Likelihood Methode - theta**



**CIR Maximum Likelihood Methode - sigma**



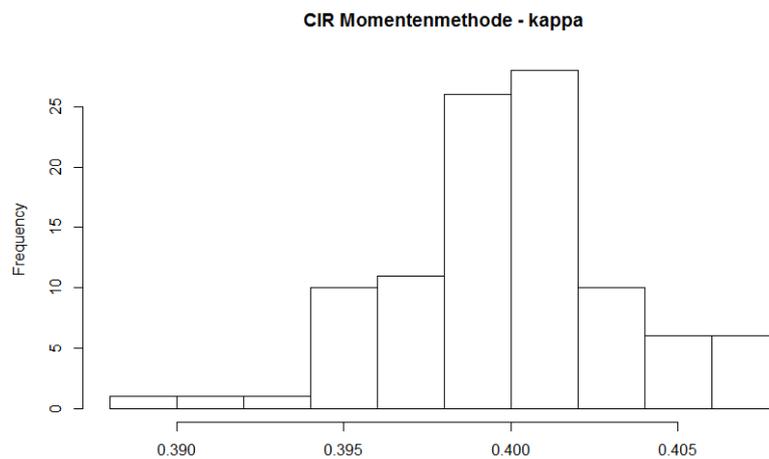
## Momentenmethode

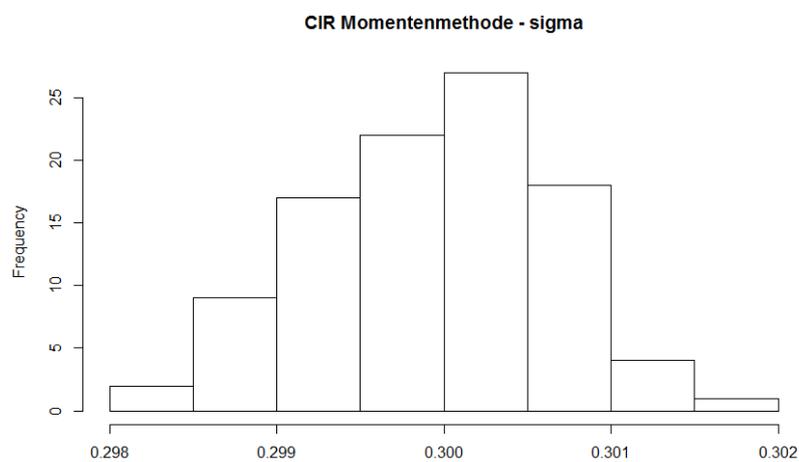
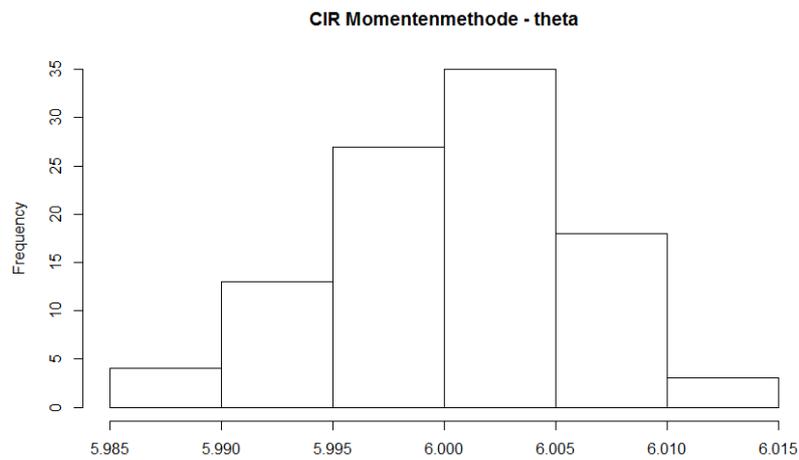
Man bemerkt bei dieser Methode ebenfalls, dass die Einträge der Varianz-Kovarianz Matrix klein sind. Bei Betrachtung der Histogramme sieht man, dass die Schätzer der Parameter sehr gut sind. Ein weiterer positiver Aspekt der Methode ist zudem der geringe Rechenaufwand.

Der theoretische Varianzschätzer, berechnet mit der Methode aus Kapitel 4.2, ist durch folgende Matrix gegeben:

$$\begin{bmatrix} 0.002732993 & 0.09569274 & -0.06768842 \\ 0.095692740 & 3.41988046 & -2.37125493 \\ -0.067688419 & -2.37125493 & 1.67647021 \end{bmatrix}$$

Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:



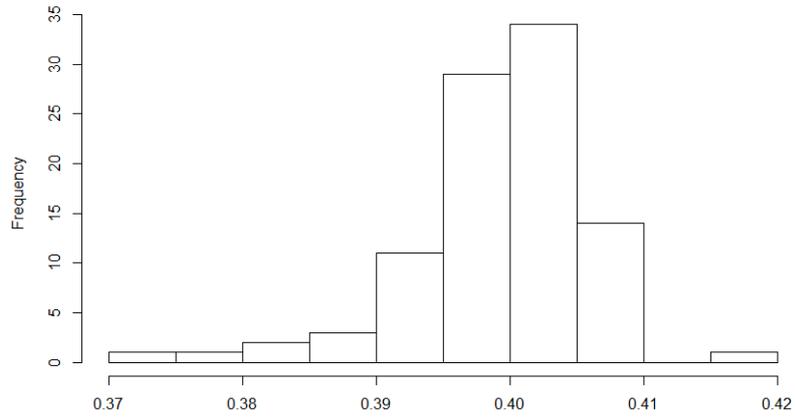


## Exponentielle Momentenmethode

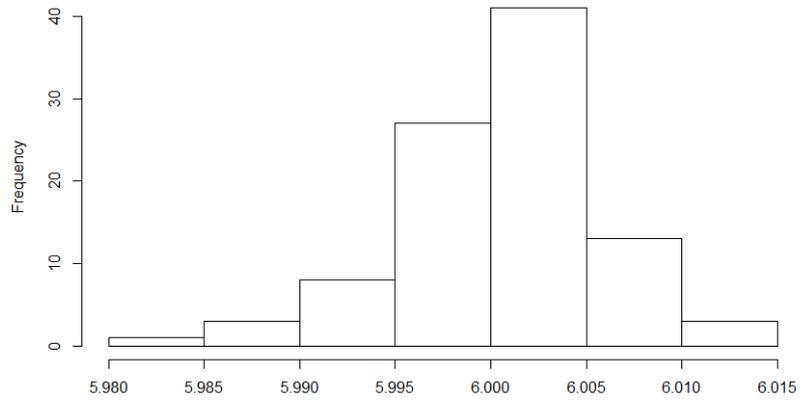
Wie bereits angemerkt können wir hier wegen der Änderung der Methode keinen theoretischen Varianzschätzer angeben. Die Histogramme zeigen uns jedoch, dass auch diese Methode sehr gute Schätzer für die Parameter liefert und die Abweichungen vom wahren Wert klein sind.

Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:

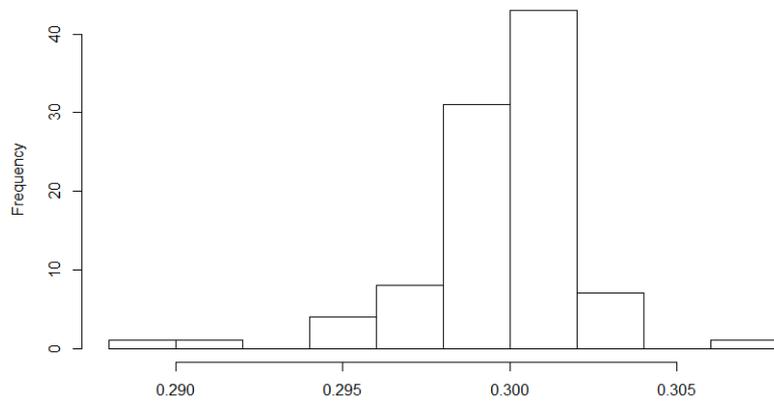
**CIR exp Momentenmethode - kappa**



**CIR exp Momentenmethode - theta**



**CIR exp Momentenmethode - sigma**



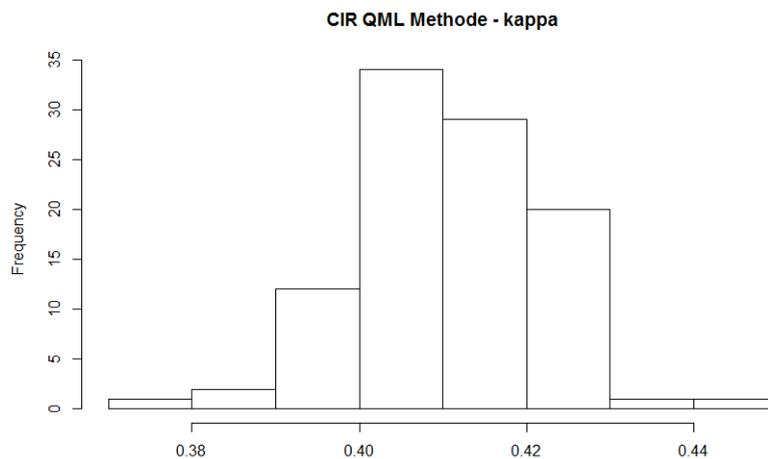
## Quasi-Maximum-Likelihood Methode

Man sieht bei dieser Methode ebenfalls sehr kleine Einträge in der Varianz-Kovarianz Matrix. Im deutlichen Unterschied zu den anderen Methoden merkt man bei Betrachtung der Histogramme hier jedoch, dass der Parameter  $\sigma$  konsequent überschätzt wird. Der eigentliche Wert  $\sigma = 0.3$  wird nicht getroffen, der Parameterschätzer bewegt sich zwischen 0.38 und 0.4. Bei den anderen beiden Parametern ist die Schätzung im Gegenzug sehr gut, vergleichbar mit den anderen Methoden.

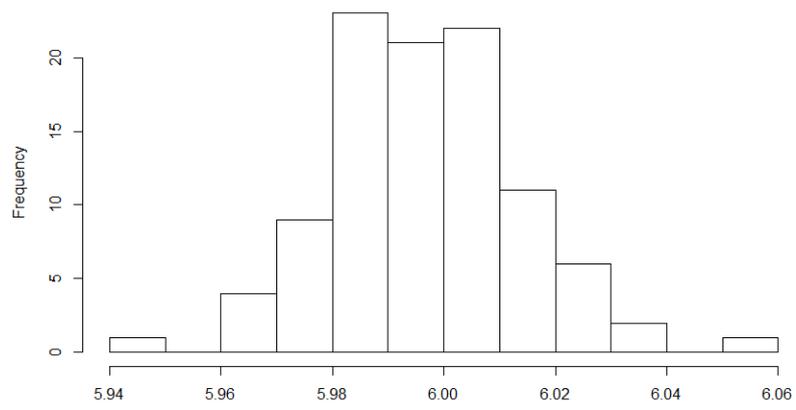
Der theoretische Varianzschätzer, berechnet mit der Methode aus Kapitel 4.4, ist durch folgende Matrix gegeben:

$$\begin{bmatrix} 1.2335668 & -0.19882196 & 0.15515616 \\ -0.1988220 & 3.46504352 & -0.06598029 \\ 0.1551562 & -0.06598029 & 0.03469615 \end{bmatrix}$$

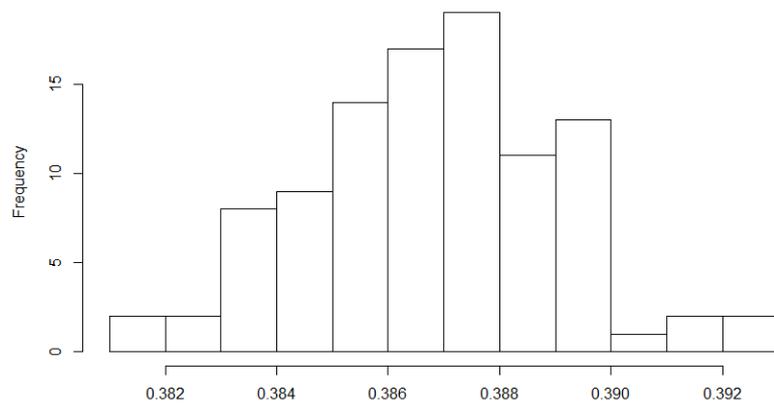
Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:



**CIR QML Methode - theta**



**CIR QML Methode - sigma**



## 6 Gamma-OU Modell

Als zweites Modell betrachten wir ein Nicht-Gaußsches Ornstein Uhlenbeck (Non-Gaussian OU) Modell, nämlich das Gamma-OU Modell. Allgemein ist ein OU-Prozess definiert als die Lösung der folgenden stochastischen Differentialgleichung:

$$dy(t) = -\lambda y(t)dt + dZ(t)$$

Hierbei ist  $Z$  ein Lévy Prozess. Wir betrachten nun im Speziellen die auf den ersten Blick seltsam erscheinende Zeitwahl für den Lévy Prozess:  $Z(\lambda t) =: Z_\lambda(t)$ , womit später die stationäre Verteilung unabhängig vom Parameter  $\lambda$  ist. Da unser Interesse in der Modellierung der Volatilität liegt, bezeichnen wir den Prozess im Folgenden mit  $V(t)$ . Somit ergibt sich:

$$dV(t) = -\lambda V(t)dt + dZ_\lambda(t), \quad V(0) = V_0 \quad (28)$$

Dabei gilt  $\lambda > 0$  und dass  $Z$  ein Subordinator, also ein wachsender Lévy Prozess, ist.

Bereits anhand dieser Definition sieht man, dass der Prozess  $V$  sich also ausschließlich durch die Sprünge des Lévy Prozesses nach oben bewegt und, bedingt durch die Struktur des Prozesses und das positive  $\lambda$ , zwischen den Sprüngen exponentiell abfällt.

Vorteile bei der Beschreibung der Volatilität durch den Nicht-Gaußschen OU Prozess, sind die folgenden (siehe [2]):

- Bei nicht negativem Startwert und der Tatsache, dass  $Z$  ein Subordinator ist, bleibt der Prozess fast sicher positiv.
- Im Falle, dass der Prozess Null erreicht, kann er daraufhin wieder positiv werden.
- Der Prozess lässt Sprünge zu, ist also nicht stetig.
- Es gibt eine stationäre Verteilung.

Als stochastischen Term verwendet der Nicht-Gaußsche OU Prozess einen wachsenden Lévy Prozess  $Z$ .

In unserem Fall, dem  $\Gamma$ -OU Modell, ist, wie der Name bereits verrät, die stationäre Verteilung eine Gamma Verteilung. Für (28) gilt, dass  $V_0 \sim \Gamma(\nu, \alpha)$

mit Parametern  $\nu > 0$  und  $\alpha > 0$ . Für den Subordinator  $Z$  gilt im  $\Gamma$ -OU Modell, dass dieser ein Compound Poisson Prozess mit Intensität  $\lambda$  und Sprunghöhenverteilung  $Exp(\alpha)$  ist. Der Compound Poisson Prozess ist ein reiner Sprungprozess.

Aufgrund seiner Funktion wird der Prozess  $Z$  in diesem Zusammenhang auch als "Background Driving Lévy Prozess" (BDLP) bezeichnet.

## 6.1 Bedingte Verteilung

Für das weitere Arbeiten mit dem  $\Gamma$ -OU Modell benötigen wir folgendes Wissen über die Verteilung:

**Satz 6.1.** *Mit  $s < t$  und  $\Delta = t - s$  gilt:*

$$\begin{aligned} \mathbb{E}[\exp(iyV_t)|V_s = v] &= \exp(iyv \exp(-\lambda\Delta)) + \nu \log(\alpha - iy \exp(\lambda\Delta)) \\ &\quad - \nu \log(\alpha - iy) \end{aligned} \quad (29)$$

*Beweis.* Um die bedingte Verteilung von  $V$  zu beschreiben, betrachten wir die charakteristische Funktion des Prozesses:

Während der stetige Teil uns keine Probleme bereitet, benötigen wir für den Sprungteil den so genannten charakteristischen Exponenten von  $Z$ :

**Definition.** *Für einen Lévy Prozess  $(Z_t)_{t \geq 0}$  existiert eine stetige Funktion  $\psi : \mathbb{R} \rightarrow \mathbb{R}$ , genannt der charakteristische Exponent von  $Z$ , sodass*

$$\mathbb{E}[\exp(iyZ(t))] = \exp(t\psi(y)) \quad (30)$$

Für unsere weiteren Überlegungen genügt es  $Z(1)$  zu betrachten, da für  $t = 1$  gilt, dass  $\psi$  die kumulantenenerzeugende Funktion von  $Z(1)$  ist und die kumulantenenerzeugende Funktion für  $Z(t)$  sich sofort aus dieser ergibt durch Hinzumultiplizieren von  $t$ . (siehe [4])

Wir betrachten zuerst die momentenerzeugende Funktion von  $Z$ . Wir wissen:  $Z(1)$  hat die Gestalt  $Z(1) = \sum_{i=1}^{N(1)} X_i$ . Die Anzahl der Sprünge ( $N(1)$ ) ist poissonverteilt mit Intensität  $\nu$  und die Sprunghöhen ( $X_i$ ) sind exponentialverteilt

mit Parameter  $\alpha$ . Somit ergibt sich für  $Z(1)$ :

$$\begin{aligned} N &\sim Poi(\nu), & M_N(u) &= \exp(\nu(\exp(u) - 1)) \\ X &\sim Exp(\alpha), & M_X(u) &= \frac{\alpha}{\alpha - u} \\ M_{Z(1)}(u) &= M_N(\log M_X(u)) = \exp\left[\nu\left(\frac{\alpha}{\alpha - u} - 1\right)\right] = \exp\left[\frac{\nu u}{\alpha - u}\right] \end{aligned} \quad (31)$$

Damit wird der charakteristische Exponent in (30) zu  $\psi(y) = \frac{\nu i y t}{\alpha - i y}$ .

Wir ziehen nun die folgende Darstellung für  $V(t)$ , die direkt aus der Definition des Prozesses ersichtlich ist (vgl. (28)), heran, wobei wieder  $s < t$  und  $\Delta = t - s$ :

$$V(t) = V(s) \exp(-\lambda\Delta) + \int_s^t \exp(-\lambda(t - u)) dZ_\lambda(u) \quad (32)$$

Zudem verwenden wir das folgende Resultat, das dem Lemma 15.1 aus [4] entspricht, das für  $Z$  Lévy Prozess und  $f : [s, T] \rightarrow \mathbb{R}$  linksstetig gilt:

$$\mathbb{E} \left[ \exp \left( i \int_s^T f(t) dZ_t \right) \right] = \exp \left( \int_s^T \psi(f(t)) dt \right)$$

wobei  $\psi(u)$  der charakteristische Exponent von  $Z$  ist. Zusammen ergibt das für die bedingte charakteristische Funktion also:

$$\mathbb{E}[\exp(iyV_t) | V_s = v] = \exp(iyv \exp(-\lambda\Delta)) + \int_s^t \psi(y \exp(-\lambda(t - u))) du \quad (33)$$

Im letzten Schritt verwenden wir die hergeleitete Darstellung des charakteristischen Exponenten, um das Integral auszurechnen.

Man beachte hierbei, dass wir in unseren Überlegungen zum charakteristischen Exponenten den Prozess  $(Z(t))_t$  betrachtet haben, wir aber eigentlich mit dem Prozess  $(Z_\lambda(t))_t = (Z(\lambda t))_t$  arbeiten. Das bedeutet insbesondere für den charakteristischen Exponenten einen extra Faktor  $\lambda$ , also  $\psi(y) = \frac{\lambda \nu i y t}{\alpha - i y}$ . Es ergibt sich damit:

$$\begin{aligned} \int_s^t \psi(y \exp(-\lambda(t - u))) du &= \int_0^\Delta \psi(y \exp(-\lambda u)) du = \int_0^\Delta \frac{\lambda \nu i y \exp(-\lambda u)}{\alpha - i y \exp(-\lambda u)} du \\ &= \nu \log(\alpha - i y \exp(-\lambda\Delta)) - \nu \log(\alpha - i y) \end{aligned}$$

Hierbei ist wie gehabt  $\Delta = t - s$ . Durch Einsetzen in (33) erhält man die behauptete Darstellung.  $\square$

Für spätere Berechnungen benötigen wir ebenfalls die ersten vier Momente von  $O_1 := \int_0^1 \exp(-\lambda(1-s))dZ_\lambda(s)$ , die ersten beiden bedingten Momente von  $V_1$  - insbesondere folgt aus den letzten beiden selbstverständlich sofort die später benötigte bedingte Varianz von  $V_1$  - sowie die bedingte Kovarianz:

**Satz 6.2.** *Es gilt mit der Bezeichnung  $O_1 := \int_0^1 \exp(-\lambda(1-s))dZ_\lambda(s)$  und  $\beta = \exp(-\lambda\Delta)$ :*

$$\begin{aligned}
\mathbb{E}_\gamma[O_1] &= (1 - \beta)\mathbb{E}_\gamma[V_0] \\
\mathbb{E}_\gamma[O_1^2] &= (1 - \beta^2)\mathbb{E}_\gamma[V_0^2] - 2\beta\mathbb{E}_\gamma[V_0]\mathbb{E}_\gamma[O_1] \\
\mathbb{E}_\gamma[O_1^3] &= (1 - \beta^3)\mathbb{E}_\gamma[V_0^3] - 3\beta^2\mathbb{E}_\gamma[V_0^2]\mathbb{E}_\gamma[O_1] - 3\beta\mathbb{E}_\gamma[V_0]\mathbb{E}_\gamma[O_1^2] \\
\mathbb{E}_\gamma[O_1^4] &= (1 - \beta^4)\mathbb{E}_\gamma[V_0^4] - 4\beta^3\mathbb{E}_\gamma[V_0^3]\mathbb{E}_\gamma[O_1] - 6\beta^2\mathbb{E}_\gamma[V_0^2]\mathbb{E}_\gamma[O_1^2] \\
&\quad - 4\beta\mathbb{E}_\gamma[V_0]\mathbb{E}_\gamma[O_1^3] \\
\mathbb{E}_\gamma[V_1|V_0 = v] &= \beta v + (1 - \beta)\mathbb{E}_\gamma[V_0] \\
\mathbb{E}_\gamma[V_1^2|V_0 = v] &= \beta^2 v^2 + 2\beta\mathbb{E}_\gamma[V_0]\mathbb{E}_\gamma[O_1] + \mathbb{E}_\gamma[O_1^2]
\end{aligned} \tag{34}$$

Die dafür benötigten ersten vier Momente von  $V_0$  sehen dabei folgendermaßen aus:

$$\begin{aligned}
\mathbb{E}_\gamma[V_0] &= \frac{\nu}{\alpha} \\
\mathbb{E}_\gamma[V_0^2] &= \frac{\nu(\nu + 1)}{\alpha^2} \\
\mathbb{E}_\gamma[V_0^3] &= \frac{\nu(\nu + 1)(\nu + 2)}{\alpha^3} \\
\mathbb{E}_\gamma[V_0^4] &= \frac{\nu(\nu + 1)(\nu + 2)(\nu + 3)}{\alpha^4}
\end{aligned}$$

Die bedingte Kovarianz ist gegeben durch

$$\begin{aligned}
\text{Cov}[V_1 V_2 | V_0 = v] &= \beta \mathbb{E}_\gamma[V_1^2 | V_0 = v] + (1 - \beta) \frac{\nu}{\alpha} \mathbb{E}_\gamma[V_1 | V_0 = v] \\
&\quad - \mathbb{E}_\gamma[V_1 | V_0 = v] \mathbb{E}_\gamma[V_2 | V_0 = v]
\end{aligned} \tag{35}$$

*Beweis.* Mit den Bezeichnungen:  $O_i := \int_{t_{i-1}}^{t_i} \exp(-\lambda(t_i - s))dZ_\lambda(s)$  und  $\beta := \exp(-\lambda\Delta)$  und unter der Berücksichtigung des speziellen Zeitgitters wird die

Darstellung von  $V(t_i) = V_i$  in (32) zu

$$V_i = \beta V_{i-1} + O_i \quad (36)$$

Unter Anwendung des Erwartungswertes auf beide Seiten und der Wahl von  $i = 1$  ergibt sich:

$$\mathbb{E}_\gamma[V_1] = \beta \mathbb{E}_\gamma[V_0] + \mathbb{E}_\gamma[O_1]$$

Unter Berücksichtigung der Stationarität erhalten wir schließlich

$$\mathbb{E}_\gamma[O_1] = (1 - \beta) \mathbb{E}_\gamma[V_0] \quad (37)$$

Weiters wenden wir den Erwartungswert auf die quadrierte Gleichung (36), wobei wieder  $i = 1$ ,  $V_1^2 = \beta^2 V_0^2 + 2\beta V_0 O_1 + O_1^2$ , an, um auf einen Ausdruck für  $\mathbb{E}_\gamma[O_1^2]$  zu erhalten:

$$\mathbb{E}_\gamma[V_1^2] = \beta^2 \mathbb{E}_\gamma[V_0^2] + 2\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1] + \mathbb{E}_\gamma[O_1^2]$$

Unter Berücksichtigung der Stationarität ergibt sich damit

$$\mathbb{E}_\gamma[O_1^2] = (1 - \beta^2) \mathbb{E}_\gamma[V_0^2] - 2\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1] \quad (38)$$

Für das dritte und vierte Moment ziehen wir wieder die Gleichung (36) heran und nehmen diese zur dritten bzw. vierten Potenz. Die Argumentation erfolgt analog mit Berücksichtigung der Stationarität.

Mit  $V_1^3 = \beta^3 V_0^3 + 3\beta^2 V_0^2 O_1 + 3\beta V_0 O_1^2 + O_1^3$  ergibt sich

$$\begin{aligned} \mathbb{E}_\gamma[V_1^3] &= \beta^3 \mathbb{E}_\gamma[V_0^3] + 3\beta^2 \mathbb{E}_\gamma[V_0^2] \mathbb{E}_\gamma[O_1] + 3\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1^2] + \mathbb{E}_\gamma[O_1^3] \\ \mathbb{E}_\gamma[O_1^3] &= (1 - \beta^3) \mathbb{E}_\gamma[V_0^3] - 3\beta^2 \mathbb{E}_\gamma[V_0^2] \mathbb{E}_\gamma[O_1] - 3\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1^2] \end{aligned}$$

Mit  $V_1^4 = \beta^4 V_0^4 + 4\beta^3 V_0^3 O_1 + 6\beta^2 V_0^2 O_1^2 + 4\beta V_0 O_1^3 + O_1^4$  ergibt sich

$$\begin{aligned} \mathbb{E}_\gamma[V_1^4] &= \beta^4 \mathbb{E}_\gamma[V_0^4] + 4\beta^3 \mathbb{E}_\gamma[V_0^3] \mathbb{E}_\gamma[O_1] + 6\beta^2 \mathbb{E}_\gamma[V_0^2] \mathbb{E}_\gamma[O_1^2] \\ &\quad + 4\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1^3] + \mathbb{E}_\gamma[O_1^4] \\ \mathbb{E}_\gamma[O_1^4] &= (1 - \beta^4) \mathbb{E}_\gamma[V_0^4] - 4\beta^3 \mathbb{E}_\gamma[V_0^3] \mathbb{E}_\gamma[O_1] - 6\beta^2 \mathbb{E}_\gamma[V_0^2] \mathbb{E}_\gamma[O_1^2] \\ &\quad - 4\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1^3] \end{aligned}$$

Nun zur eigentlichen Berechnung der Erwartungswerte. Wir verwenden dabei, dass  $O_1$  unabhängig von  $V_0$  ist sowie (37) und (38).

$$\begin{aligned}
\mathbb{E}_\gamma[V_1|V_0 = v] &= \mathbb{E}_\gamma[\beta V_0 + O_1|V_0 = v] = \beta v + \mathbb{E}_\gamma[O_1] \\
&= \beta v + (1 - \beta)\mathbb{E}_\gamma[V_0] \\
\mathbb{E}_\gamma[V_1^2|V_0 = v] &= \mathbb{E}_\gamma[\beta^2 V_0^2 + 2\beta V_0 O_1 + O_1^2|V_0 = v] \\
&= \beta^2 v^2 + 2\beta \mathbb{E}_\gamma[V_0] \mathbb{E}_\gamma[O_1] + \mathbb{E}_\gamma[O_1^2]
\end{aligned} \tag{39}$$

Um die benötigten Momente von  $V_0$  auszurechnen, erinnern wir uns, dass  $V_0$  gammaverteilt ist mit Parametern  $\nu$  und  $\alpha$ . Die angegebenen Momente erhält man dann einfach mithilfe der momentenerzeugenden Funktion der Gamma Verteilung, nämlich:

$$M(t; \nu, \alpha) = \left( \frac{\alpha}{\alpha - t} \right)^\nu$$

Für die Darstellung der bedingten Kovarianz stellen wir zunächst folgende Überlegungen an, wobei wir Eigenschaften der bedingten Erwartung verwenden:

$$\begin{aligned}
\mathbb{E}_\gamma[V_1 V_2 | V_0] &= \mathbb{E}_\gamma[\mathbb{E}_\gamma[V_1 V_2 | V_1] | V_0] = \mathbb{E}_\gamma[V_1 \mathbb{E}_\gamma[V_2 | V_1] | V_0] \\
&= \mathbb{E}_\gamma[V_1 (\beta V_1 + (1 - \beta) \frac{\nu}{\alpha}) | V_0] \\
&= \beta \mathbb{E}_\gamma[V_1^2 | V_0] + (1 - \beta) \frac{\nu}{\alpha} \mathbb{E}_\gamma[V_1 | V_0]
\end{aligned}$$

Damit folgt die behauptete Darstellung für die bedingte Kovarianz:

$$\begin{aligned}
\text{Cov}[V_1 V_2 | V_0] &= \mathbb{E}_\gamma[V_1 V_2 | V_0] - \mathbb{E}_\gamma[V_1 | V_0] \mathbb{E}_\gamma[V_2 | V_0] \\
&= \beta \mathbb{E}_\gamma[V_1^2 | V_0] + (1 - \beta) \frac{\nu}{\alpha} \mathbb{E}_\gamma[V_1 | V_0] - \mathbb{E}_\gamma[V_1 | V_0] \mathbb{E}_\gamma[V_2 | V_0]
\end{aligned}$$

wobei man  $\mathbb{E}_\gamma[V_2 | V_0]$  durch analoge Argumentation und Rechnung wie vorher erhält als:

$$\mathbb{E}_\gamma[V_2 | V_0 = v] = \mathbb{E}_\gamma[\beta^2 V_0 + O_2 | V_0 = v] = \beta^2 v + \mathbb{E}_\gamma[O_2] = \beta^2 v + (1 - \beta^2) \mathbb{E}_\gamma[V_0]$$

□

## 6.2 Stationäre Verteilung

Wie bereits anfangs erwähnt, besitzt der  $\Gamma$ -OU Prozess eine stationäre Verteilung, nämlich eine Gamma Verteilung mit Parametern  $\nu$  und  $\alpha$ . Ausgehend davon kann man bereits zeigen, dass der BDLP  $Z$  wie behauptet einer Compound Poisson Verteilung mit Intensität  $\nu$  und exponentialverteilten Sprüngen mit Parameter  $\alpha$  folgen muss.

Wir verwenden dafür die folgende Fundamentalgleichung, die Lemma 3.1 aus [6] entspricht. Diese liefert eine Aussage über den Zusammenhang der kumulantenerzeugenden Funktion der stationären Verteilung, im folgenden bezeichnet als  $K_D(u; \nu, \alpha)$ , und der kumulantenerzeugenden Funktion der Verteilung von  $Z(1)$ , im folgenden bezeichnet als  $K_Z(u; \nu, \alpha)$ :

$$K_Z(u; \nu, \alpha) = u \frac{\partial}{\partial u} K_D(u; \nu, \alpha) \quad (40)$$

Die kumulantenerzeugende Funktion der Gamma Verteilung sieht folgendermaßen aus:

$$K_D(u; \nu, \alpha) = \log(M_D(u; \nu, \alpha)) = \log\left(\left(\frac{\alpha}{\alpha - u}\right)^\nu\right) = \nu \log\left(\frac{\alpha}{\alpha - u}\right)$$

Ableiten nach  $u$  und Einsetzen in (40) ergibt schließlich die kumulantenerzeugende Funktion der Verteilung von  $Z(1)$ :

$$K_Z(u; \nu, \alpha) = u \frac{\partial}{\partial u} K_D(u; \nu, \alpha) = u \frac{\nu}{\alpha - u}$$

Da eine Verteilung eindeutig durch ihre kumulantenerzeugende Funktion charakterisiert ist, sind wir hiermit bereits fertig, denn das Ergebnis von  $K_Z$  ist die kumulantenerzeugende Funktion eines Compound Poisson Prozesses mit Intensität  $\nu$  und exponentialverteilten Sprüngen mit Parameter  $\alpha$  (vgl. momentenerzeugende Funktion davon in (31)).

## 6.3 Simulation

Um Beobachtungen simulieren zu können, brauchen wir im Wesentlichen die Übergänge von  $V$  von einem Zeitschritt auf den nächsten auf dem Zeitgitter

$0 = t_0 < t_1 < \dots < t_N$ . Für beliebiges  $i \in [1, N]$  also:

$$V(t_i) = V(t_{i-1}) \exp(-\lambda(t_i - t_{i-1})) + \int_{t_{i-1}}^{t_i} \exp(-\lambda(t_i - s)) dZ_\lambda(s) \quad (41)$$

Für das Integral simulieren wir zunächst für diesen einen Zeitschritt den Compound Poisson Prozess  $Z_\lambda$ , wobei wir folgende Tatsachen verwenden, siehe [4]:

- Die Anzahl der Sprünge  $N(T)$  in einem Intervall  $[0, T]$  eines Compound Poisson Prozesses ist poissonverteilt mit Parameter  $\nu T$ , wobei  $\nu$  die Intensität des Compound Poisson Prozesses ist.
- Die genauen Sprungzeiten sind, bedingt auf die Anzahl der Sprünge  $N(T)$ , unabhängig voneinander und auf dem Intervall  $[0, T]$  gleichverteilt.

Die Simulation des Compound Poisson Prozesses  $Z_\lambda$  geschieht also mit folgenden Schritten:

1. Simulation der Anzahl der Sprünge  $N(T)$ . Dies ist eine poissonverteilte Zufallsvariable mit Parameter  $\nu T$ , wobei in unserem Fall, da wir  $Z_\lambda(t) = Z(\lambda t)$  betrachten, noch mit  $\lambda$  multipliziert werden muss, also  $\nu \lambda T$ .
2. Simulation der Sprungzeiten  $U_i$  mit  $N(T)$  unabhängigen, gleichverteilten Zufallsvariablen auf dem Intervall  $[0, T]$ .
3. Simulation der Sprunghöhen  $Y_i$  mit  $N(T)$  unabhängigen Zufallsvariablen, die im allgemeinen Fall exponentialverteilt sind mit Parameter  $\frac{\alpha}{\lambda}$ . Bei unserer Betrachtung von  $Z_\lambda$  wird dies zu  $\lambda \frac{\alpha}{\lambda} = \alpha$ .

Der eigentliche Compound Poisson Prozess setzt sich dann folgendermaßen zusammen:

$$Z_\lambda(t) = \sum_{i=1}^{N(T)} \mathbb{1}_{(U_i < t)} Y_i$$

Nun, mithilfe der Sprungzeiten  $U_1, \dots, U_{N(T)}$  und den dazugehörigen Sprunghöhen  $Y_1, \dots, Y_{N(T)}$  ergibt sich für das Integral:

$$\int_{t_{i-1}}^{t_i} \exp(-\lambda(t_i - s)) dZ_\lambda(s) = \sum_{j=1}^{N(1)} \exp(-\lambda(t_i - U_j)) Y_j$$

Im Folgenden wählen wir zur Vereinfachung als Zeitgitter jene äquidistanten Gitterpunkte, die zueinander Abstand  $\Delta$  haben, also  $t_j = t_{j-1} + \Delta$ . Weiters

schreiben wir:  $V_i := V(t_i)$

Der folgende R-Code implementiert die beschriebene Simulation des Prozesses  $V$ :

---

Simulation OU

---

```
nu=53.3334
```

```
al=8.8889
```

```
la=0.4
```

```
N=1000
```

```
dt=1
```

```
V<-vector(length=N)
```

```
V[1]=rgamma(1,nu,al)
```

```
for(j in 2:N){
```

```
  Nt <- rpois(1,nu*la*dt)
```

```
  U <- runif(Nt,0,dt)
```

```
  Y <- rexp(Nt,al)
```

```
  V[j]=V[j-1]*exp(-la*dt)+sum(exp(-la*(dt-U))*Y)
```

```
}
```

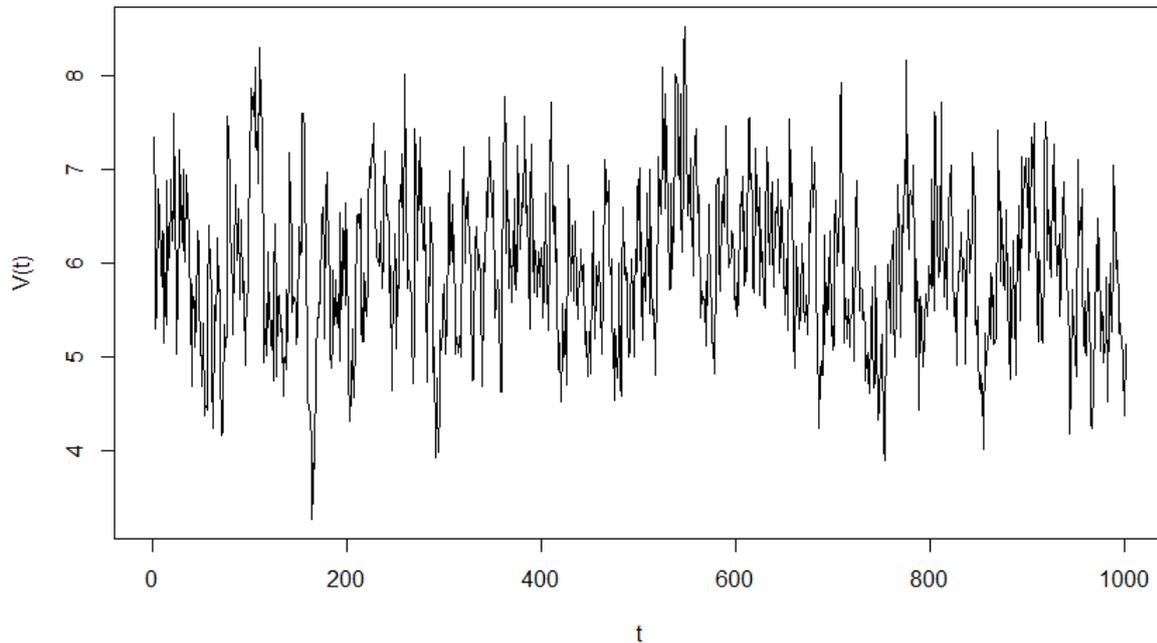
---

Hier ein Beispiel für eine solche Simulation mit

$$\nu = 53.3334, \alpha = 8.8889, \lambda = 0.4, \Delta = 1, N = 1000.$$

Die Parameter wurden dabei so gewählt, dass der Prozess das gleiche stationäre Mittel und die gleiche stationäre Varianz wie der Prozess im Beispiel beim CIR Prozess hat.

### Simulation Gamma-OU



## 7 Parameterschätzung Gamma-OU Modell

Nach dieser Vorarbeit können wir nun mit der eigentlichen Parameterschätzung beginnen. Dafür vergleichen wir wieder folgende Methoden:

- Momentenmethode
- Exponentielle Momentenmethode
- Quasi-Maximum-Likelihood Methode

Im Unterschied zum CIR Modell betrachten wir die Maximum-Likelihood Methode nicht. Wie in [8] bemerkt, kann im  $\Gamma$ -OU Fall die Maximum-Likelihood Methode nicht in ihrer üblichen Weise definiert werden.

Das Problem ist, dass der Prozess  $V$  von einem Compound Poisson Prozess angetrieben wird, zwischen den Sprüngen exponentiell abklingt, und somit die Übergangsverteilung ein Atom (für den Fall dass kein Sprung passiert) besitzt, dessen Position vom kontinuierlichen Parameter  $\lambda$  abhängt. Deswegen

besitzt das statistische Experiment kein dominierendes  $\sigma$ -endliches Maß, was üblicherweise in der Definition der Dichten und Likelihood-Funktion vorausgesetzt wird.

Insgesamt gibt es beim  $\Gamma$ -OU Modell die drei Parameter:  $\lambda, \nu$  und  $\alpha$ . Alternativ kann man statt der Parameter der Gamma Verteilung,  $\nu$  und  $\alpha$ , auch den Erwartungswert  $\zeta = \frac{\nu}{\alpha}$  und die Varianz  $\eta = \frac{\nu}{\alpha^2}$  heranziehen. Beide Parametrisierungen können offensichtlich aus der jeweilig anderen gewonnen werden und beide haben ihre Vorteile. Während es bisher einfacher war die erste Parametrisierung zu verwenden, hat die zweite Parametrisierung mit  $\gamma = (\lambda, \zeta, \eta)$  besonders bei der Varianzschätzung Vorteile, wie wir später sehen werden.

Es ist also im Folgenden der Parameterraum  $\Gamma = \{\gamma = (\lambda, \nu, \alpha) \in \mathbb{R}^3 : \lambda, \nu, \alpha > 0\}$  bzw. später  $\Gamma = \{\gamma = (\lambda, \zeta, \eta) \in \mathbb{R}^3 : \lambda, \zeta, \eta > 0\}$  und der Parametervektor mit den wahren Parametern sei  $\gamma_0$ . Die  $N$  Beobachtungen, die unter  $\gamma_0$  simuliert wurden, werden mit  $V_1, \dots, V_N$  bezeichnet. Wir werden viele Schätzer iterativ aus entsprechenden Gleichungssystemen ermitteln. Für die Startwerte für diese Iterationen ziehen wir die folgenden einfachen unbedingten Momentenschätzer heran, die sich mithilfe der Formeln aus (34), (35) ergeben, wobei  $\beta = \exp(-\lambda\Delta)$ :

$$\begin{aligned}\mathbb{E}[V_1] &= \beta\mathbb{E}_\gamma[V_0] + (1 - \beta)\mathbb{E}_\gamma[V_0] = \mathbb{E}_\gamma[V_0] = \zeta \\ \mathbb{V}[V_1] &= \beta^2\mathbb{E}_\gamma[V_0^2] + (1 - \beta^2)\mathbb{E}_\gamma[V_0^2] - \zeta^2 = \zeta^2 + \eta - \zeta^2 = \eta \\ \text{Cov}[V_1V_2] &= \beta(\zeta^2 + \eta) + \zeta(1 - \beta)\zeta - \zeta^2 \\ &= \beta\eta \Rightarrow \rho(1) = \frac{\text{Cov}[V_1V_2]}{\sqrt{\mathbb{V}[V_1]}\sqrt{\mathbb{V}[V_2]}} = \beta\end{aligned}$$

Somit sind  $\lambda, \zeta$  und  $\eta$ :

$$\begin{aligned}\lambda &= -\frac{1}{\Delta} \log(\hat{\rho}(1)), \quad \text{wobei } \hat{\rho}(1) = \frac{\sum_{j=1}^{N-1} (V_{j+1} - \zeta)(V_j - \zeta)}{\sum_{j=1}^N (V_j - \zeta)^2} \\ \zeta &= \frac{1}{N} \sum_{j=1}^N V_j \\ \eta &= \frac{1}{N} \sum_{j=1}^N (V_j - \zeta)^2\end{aligned}\tag{42}$$

Startwerte für  $\nu$  und  $\alpha$  ergeben sich durch  $\nu = \frac{\zeta^2}{\eta}$  bzw.  $\alpha = \frac{\zeta}{\eta}$ .

## 7.1 Momentenmethode

Wir konstruieren die Martingalschätzer wie beim vorigen Modell in der folgenden Art:

$$\begin{aligned}
G_N^1(\gamma) &= \sum_{j=1}^N [V_j - f^1(V_{j-1}, \gamma)], & f^1(v, \gamma) &= \mathbb{E}_\gamma[V_1|V_0 = v] \\
G_N^2(\gamma) &= \sum_{j=1}^N [V_j V_{j-1} - f^2(V_{j-1}, \gamma)], & f^2(v, \gamma) &= \mathbb{E}_\gamma[V_1 V_0|V_0 = v] \\
G_N^3(\gamma) &= \sum_{j=1}^N [V_j^2 - f^3(V_{j-1}, \gamma)], & f^3(v, \gamma) &= \mathbb{E}_\gamma[V_1^2|V_0 = v]
\end{aligned} \tag{43}$$

Hierbei ist  $\mathbb{E}_\gamma$  der Erwartungswert bezüglich des Parametervektors  $\gamma = (\lambda, \nu, \alpha)$ .

Nun zur eigentlichen Berechnung der Erwartungswerte in (43). Wir verwenden dabei, dass  $O_1$  unabhängig von  $V_0$  ist sowie die Ergebnisse aus (34), insbesondere die beiden Gleichungen (37) und (38). Wie gehabt ist dabei  $\beta = \exp(-\lambda\Delta)$ .

$$\begin{aligned}
f^1(v, \gamma) &= \mathbb{E}_\gamma[V_1|V_0 = v] = \mathbb{E}_\gamma[V_0\beta + O_1|V_0 = v] = v\beta + \mathbb{E}_\gamma[O_1] \\
&= v\beta + (1 - \beta)\mathbb{E}_\gamma[V_0] = v\beta + (1 - \beta)\frac{\nu}{\alpha} \\
f^2(v, \gamma) &= \mathbb{E}_\gamma[V_1 V_0|V_0 = v] = v^2\beta + v(1 - \beta)\frac{\nu}{\alpha} \\
f^3(v, \gamma) &= \mathbb{E}_\gamma[V_1^2|V_0 = v] = \mathbb{E}_\gamma[\beta^2 V_0^2 + 2\beta V_0 O_1 + O_1^2|V_0 = v] \\
&= \beta^2 v^2 + 2\beta v(1 - \beta)\frac{\nu}{\alpha} + (1 - \beta^2)\mathbb{E}_\gamma[V_0^2] - 2\beta\mathbb{E}_\gamma[V_0]\mathbb{E}_\gamma[O_1] \\
&= \beta^2 v^2 + 2\beta v(1 - \beta)\frac{\nu}{\alpha} + (1 - \beta^2)\frac{\nu(\nu + 1)}{\alpha^2} - 2\beta\frac{\nu}{\alpha}(1 - \beta)\frac{\nu}{\alpha} \\
&= \beta^2 v^2 + 2\beta v(1 - \beta)\frac{\nu}{\alpha} + (1 - \beta^2)\frac{\nu}{\alpha^2} + ((1 - \beta^2) - 2\beta(1 - \beta))\frac{\nu^2}{\alpha^2} \\
&= \beta^2 v^2 + 2\beta v(1 - \beta)\frac{\nu}{\alpha} + (1 - \beta)^2\frac{\nu^2}{\alpha^2} + (1 - \beta^2)\frac{\nu}{\alpha^2}
\end{aligned} \tag{44}$$

Den Schätzer  $\hat{\gamma}_N$  des Parametervektors  $\gamma = (\lambda, \nu, \theta)$  erhalten wir, indem wir die drei Gleichungen (43) gleich Null setzen, also  $G_N^j = 0, j = 1, 2, 3$ . Wir können diese drei Gleichungen mit drei Unbekannten, den Parametern, explizit lösen. Damit dies übersichtlich bleibt, führen wir die folgenden Bezeichnungen ein:

$$\begin{aligned}\zeta &:= \frac{\nu}{\alpha}, & \eta &:= \frac{\nu}{\alpha^2}, & \beta &:= \exp(-\lambda\Delta) \\ \xi_{1,N} &:= \frac{1}{N} \sum_{i=1}^N V_i, & \xi_{2,N} &:= \frac{1}{N} \sum_{i=1}^N V_i V_{i-1} & \xi_{3,N} &:= \frac{1}{N} \sum_{i=1}^N V_i^2 \\ v_{1,N} &= \frac{1}{N} \sum_{i=1}^N V_{i-1}, & v_{2,N} &= \frac{1}{N} \sum_{i=1}^N V_{i-1}^2\end{aligned}$$

Setzen wir nun unsere Erkenntnisse aus (44) in die Gleichungen  $G_N^i, i = 1, 2, 3$  ein, erhalten wir:

$$\begin{aligned}G_N^1(\gamma) &= \sum_{j=1}^N [V_j - f^1(V_{j-1}, \gamma)] = \sum_{j=1}^N [V_j - V_{j-1}\beta - (1 - \beta)\zeta] \\ &= \sum_{j=1}^N V_j - \beta \sum_{j=1}^N V_{j-1} - N(1 - \beta)\zeta = N\xi_1 - N\beta v_1 - N(1 - \beta)\zeta \\ &= N[\xi_1 - \beta v_1 - (1 - \beta)\zeta] \stackrel{!}{=} 0 \\ G_N^2(\gamma) &= \sum_{j=1}^N [V_j V_{j-1} - f^2(V_{j-1}, \gamma)] = \sum_{j=1}^N [V_j V_{j-1} - V_{j-1}^2\beta - V_{j-1}(1 - \beta)\zeta] \\ &= \sum_{j=1}^N V_j V_{j-1} - \beta \sum_{j=1}^N V_{j-1}^2 - (1 - \beta)\zeta \sum_{j=1}^N V_{j-1} \\ &= N\xi_{2,N} - N\beta\xi_{3,N} - (1 - \beta)\zeta N v_{1,N} \\ &= N[\xi_{2,N} - \beta\xi_{3,N} - (1 - \beta)\zeta v_{1,N}] \stackrel{!}{=} 0 \\ G_N^3(\gamma) &= \sum_{j=1}^N [V_j^2 - f^3(V_{j-1}, \gamma)] \\ &= \sum_{j=1}^N [V_j^2 - \beta^2 V_{j-1}^2 - 2\beta V_{j-1}(1 - \beta)\zeta - (1 - \beta)^2 \zeta^2 - (1 - \beta^2)\eta]\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^N V_j^2 - \beta^2 \sum_{j=1}^N V_{j-1}^2 - 2\beta(1-\beta)\zeta \sum_{j=1}^N V_{j-1} - N(1-\beta)^2\zeta^2 - N(1-\beta^2)\eta \\
&= N\xi_{3,N} - N\beta^2 v_{2,N} - 2\beta(1-\beta)\zeta N v_{1,N} - N(1-\beta)^2\zeta^2 - N(1-\beta^2)\eta \\
&= N[\xi_{3,N} - \beta^2 v_{2,N} - 2\beta(1-\beta)\zeta v_{1,N} - (1-\beta)^2\zeta^2 - (1-\beta^2)\eta] \stackrel{!}{=} 0
\end{aligned}$$

Wir lösen nun diese drei Gleichungen für die Unbekannten  $\beta, \zeta$  und  $\eta$  und erhalten nach einfacher Rechnung:

$$\begin{aligned}
\beta &= \frac{\xi_{2,N} - \xi_{1,N} v_{1,N}}{\xi_{3,N} + v_{1,N}^2}, & \zeta &= \frac{\xi_{1,N} - \beta v_{1,N}}{1 - \beta}, \\
\eta &= \frac{\xi_{3,N} - \xi_{1,N}^2 - \beta^2(v_{2,N} - v_{1,N}^2)}{1 - \beta^2}
\end{aligned} \tag{45}$$

Dadurch erhalten wir auch sofort die Schätzer für unsere drei Parameter  $\lambda, \nu$  und  $\alpha$ , da diese ja im folgenden Zusammenhang zu den eben berechneten stehen:

$$\hat{\lambda} = -\frac{1}{\Delta} \log(\beta), \quad \hat{\nu} = \frac{\zeta^2}{\eta}, \quad \hat{\alpha} = \frac{\zeta}{\eta}$$

Die Implementierung in R sieht folgendermaßen aus:

---

#### Momentenmethode OU

---

```

xi1 <- mean(V[2:N])
xi2 <- mean(V[2:N]*V[1:(N-1)])
xi3 <- mean(V[2:N]^2)
v1 <- mean(V[1:(N-1)])
v2 <- mean(V[1:(N-1)]^2)
be <- (xi2 - xi1*v1)/(v2-v1^2)
ze <- (xi1 - be*v1)/(1-be)
eta <- ((xi3 - xi1^2) - be^2*(v2 - v1^2))/(1-be^2)

la_est <- -1/dt log(be)
al_est <- ze/eta
nu_est <- ze*al_est

```

---

## 7.2 Exponentielle Momentenmethode

Mithilfe der bedingten charakteristischen Funktion konstruieren wir wie vorher Schätzer der Art:

$$G_N^k(\gamma) = \sum_{j=1}^N [\exp(iy_k V_j) - f(V_{j-1}, \gamma)], \quad k = 1, 2, 3 \quad (46)$$
$$f(v, \gamma) = \mathbb{E}_\gamma[\exp(iy_k V_1) | V_0 = v],$$

Die dafür benötigte charakteristische Funktion des Prozesses  $V$  haben wir bereits ausgerechnet in (29).

Beim Testen dieser Methode stellte sich wie schon beim CIR Modell heraus, dass die Wahl der  $y_k$  deutlich die Qualität der Schätzung beeinflusst, ebenso gravierend wie das beim vorherigen Modell der Fall war.

Aus diesem Grund wurde hier wieder der unkonventionelle Zugang von vorher umgesetzt, vgl. Kapitel 3.3. Der folgende R-Code implementiert diese Methode:

```
Exponentielle Momentenmethode OU
-----
fexp <- function(v, nu, al, la, y){
  exp(1i*y*v*exp(-la*dt)) + nu*log(al-1i*y*exp(-la*dt))-
  nu*log(al-1i*y)}

testy <- function(y1) Re(mean(exp(1i*y1*V[2:N]) -
  fexp(V[1:(N-1)], nu1, al1, la1, y1)))
testy <- Vectorize(testy)
yroot<-uniroot.all(testy, c(-2,2))

y1 <- yroot[2]
y2 <- yroot[3]
y3 <- yroot[5]

library(rootSolve)

modeexp <- function(y)
  c(F1 = Re(mean(exp(1i*y1*V[2:N]) -
```

$$\begin{aligned}
& \text{fexp}(V[1:(N-1)], y[1], y[2], y[3], y1)), \\
F2 = & \mathbf{Re}(\mathbf{mean}(\mathbf{exp}(1i*y2*V[2:N]) - \\
& \text{fexp}(V[1:(N-1)], y[1], y[2], y[3], y2))), \\
F3 = & \mathbf{Re}(\mathbf{mean}(\mathbf{exp}(1i*y3*V[2:N]) - \\
& \text{fexp}(V[1:(N-1)], y[1], y[2], y[3], y3))))
\end{aligned}$$

```
optExpMo <- multiroot(f = modeexp,
  start = c(nu1, al1, la1))
```

---

Das Lösen des Gleichungssystems wurde wieder mit der Funktion `multiroot` aus dem Paket `rootSolve` vorgenommen.

### 7.3 Quasi-Maximum-Likelihood Methode

Wie beim vorigen Modell werden hier normalverteilte Übergangsdichten angenommen. Im Gegensatz zur Maximum-Likelihood Methode können wir die Quasi-Maximum-Likelihood Methode also umsetzen. Dafür schätzen wir Erwartungswert und Varianz aus unseren simulierten Daten. Die log-likelihood Funktion sieht also wie gehabt folgendermaßen aus:

$$l_N(\gamma) = \log(\eta(\bar{\mu}_1, \bar{\sigma}_1^2; \gamma)) + \sum_{j=2}^N \log(\eta(V_j, \mu(V_{j-1}), \sigma^2(V_{j-1}); \gamma)) \quad (47)$$

wobei  $\eta$  die Dichte der Normalverteilung bezeichnet.

Beim ersten Summanden bedienen wir uns wieder den Schätzungen für Erwartungswert und Varianz geschätzt aus der stationären Verteilung, einer Gamma Verteilung mit Parametern  $\zeta = \frac{\nu}{\alpha}, \eta = \frac{\nu}{\alpha^2}$ , und für alle weiteren Summanden erfolgt die Schätzung mit dem bedingten Erwartungswert und der bedingten Varianz, welche folgen aus (34). Das bedeutet also:

$$\begin{aligned}
\bar{\mu}_1 &= \frac{\nu}{\alpha} = \zeta \\
\bar{\sigma}_1^2 &= \frac{\nu}{\alpha^2} = \eta \\
\mu(V_{j-1}) &= \mathbb{E}_\gamma[V_j | V_{j-1} = v] = \exp(-\lambda\Delta)v + (1 - \exp(-\lambda\Delta))\zeta \\
\sigma^2(V_{j-1}) &= \mathbb{V}_\gamma[V_j | V_{j-1} = v] = (1 - \exp(-2\lambda\Delta))\eta
\end{aligned} \quad (48)$$

Diese Methode wird durch den folgenden Code implementiert:

---

Quasi-Maximum-Likelihood Methode OU

---

```
a <- function(v, nu, al, la){
  exp(-la*dt)*v+(1-exp(-la*dt))*nu/al
}

b <- function(r, ka, th, si){
  (1-(exp(-la*dt))^2)*nu/al^2
}

g <- function(nu, al, la){
  mul <- nu/al
  sil <- nu/al^2
  gstart <- log(dnorm(V[1], mul, sil))
  for (j in 2:N){
    ghelp <- log(dnorm(V[j], a(V[j-1], nu, al, la),
                      b(V[j-1], nu, al, la)))
    gstart <- gstart+ghelp
  }
  return((-1)*gstart)
}

library(stats4)

optQML<-mle(g, start=list(nu=nu1, al=al1, la=la1))
```

---

Die Funktionen  $a$  und  $b$  implementieren die obigen bedingten Erwartungswerte und bedingten Varianzen  $\mu(r_{j-1})$  bzw.  $\sigma^2(r_{j-1})$ . Die Maximierung wird mit der Funktion `mle` vorgenommen.

## 8 Varianzschätzung Gamma-OU Modell

### 8.1 Varianzschätzung Momentenmethode

Wie schon beim CIR Modell verwenden wir zur theoretischen Schätzung der Varianzen der Parameter die in [8] beschriebene Vorgehensweise:

Analog wie in Kapitel 4.2 beginnen wir mit der Berechnung der Matrix  $Y$ , wobei diese sich durch folgendes Resultat ergibt (vgl. Satz 4.2):

*Es gilt*

$$\frac{1}{\sqrt{N}}G_N(\gamma_0) \xrightarrow{D} N(0, Y)$$

mit  $N \rightarrow \infty$ , wobei die Matrix  $Y$  folgendermaßen aussieht:

$$Y_{ij} = \mathbb{E}(\text{Cov}[\Xi_k^i, \Xi_k^j | V_0])$$

Hierbei ist  $\Xi_k = (V_k, V_k V_{k-1}, V_k^2)^T$ .

Explizites Ausrechnen der Matrix  $Y$  ergibt unter Verwendung von (34) und der Tatsache, dass  $O_1$  unabhängig von  $V_0$  ist:

**Matrix  $Y$**

$$Y_{11} = \mathbb{E}(\text{Cov}[V_1, V_1 | V_0]) = \mathbb{E}(\mathbb{V}[V_1 | V_0]) = \mathbb{V}[O_1]$$

$$Y_{12} = Y_{21} = \mathbb{E}(\text{Cov}[V_1, V_1 V_0 | V_0]) = \mathbb{E}[V_0] \mathbb{V}[V_1 | V_0] = \mathbb{E}[V_0] Y_{11}$$

$$Y_{13} = Y_{31} = \mathbb{E}(\text{Cov}[V_1, V_1^2 | V_0]) = \mathbb{E}(\text{Cov}[\beta V_0 + O_1, \beta^2 V_0^2 + 2\beta V_0 O_1 + O_1^2 | V_0])$$

$$= 2\beta \mathbb{E}[V_0] \mathbb{E}(\text{Cov}[O_1, O_1 | V_0]) + \mathbb{E}(\text{Cov}[O_1, O_1^2 | V_0])$$

$$= 2\beta \mathbb{E}[V_0] \mathbb{V}[O_1] + \text{Cov}[O_1, O_1^2] = \mathbb{E}[V_0] \mathbb{V}[O_1] + \mathbb{E}[O_1^3] - \mathbb{E}[O_1] \mathbb{E}[O_1^2]$$

$$Y_{22} = \mathbb{E}(\text{Cov}[V_1 V_0, V_1 V_0 | V_0]) = \mathbb{E}[V_0^2] \mathbb{E}(\mathbb{V}[V_1 | V_0]) = \mathbb{E}[V_0^2] Y_{11}$$

$$Y_{23} = Y_{32} = \mathbb{E}(\text{Cov}[V_1 V_0, V_1^2 | V_0]) = \mathbb{E}[V_0] \mathbb{E}(\text{Cov}[V_1, V_1^2 | V_0]) = \mathbb{E}[V_0] Y_{13}$$

$$Y_{33} = \mathbb{E}(\text{Cov}[V_1^2, V_1^2 | V_0]) = \mathbb{E}(\mathbb{V}[\beta^2 V_0^2 + 2\beta V_0 O_1 + O_1^2 | V_0])$$

$$= 4\beta^2 \mathbb{E}[V_0^2] \mathbb{E}(\text{Cov}[O_1, O_1 | V_0]) + 4\beta \mathbb{E}[V_0] \mathbb{E}(\text{Cov}[O_1, O_1^2 | V_0]) + \mathbb{E}(\text{Cov}[O_1^2, O_1^2 | V_0])$$

$$= 4\beta^2 \mathbb{E}[V_0^2] \mathbb{V}[O_1] + 4\beta \mathbb{E}[V_0] \text{Cov}[O_1, O_1^2] + \text{Cov}[O_1^2, O_1^2]$$

$$= 4\beta^2 \mathbb{E}[V_0^2] \mathbb{V}[O_1] + 4\beta \mathbb{E}[V_0] (\mathbb{E}[O_1^3] - \mathbb{E}[O_1] \mathbb{E}[O_1^2]) + \mathbb{E}[O_1^4] - \mathbb{E}[O_1^2]^2$$

Im zweiten Schritt gehen wir über zur Matrix  $\Sigma$ . Diese ergibt sich folgendermaßen (vgl. Satz 4.3):

Es gilt

$$\sqrt{N}[\xi_N - \xi] \xrightarrow{D} N(0, \Sigma)$$

wobei

$$\Sigma = P^{-1}Y(P^{-1})^T$$

und

$$P_{ij} = \delta_{ij} - \phi_1^i \delta_{1j} - \phi_2^i \delta_{3j}$$

Hierbei bezeichnet  $\delta_{ij}$  das Kroneker Delta und die  $\phi_l^i$  erhalten wir aus der folgenden Darstellung der Funktionen  $f^i(v, \gamma)$ ,  $i = 1 \dots 3$ :

$$f^i(v, \gamma) = \phi_2^i r^2 + \phi_1^i r + \phi_0^i$$

Explizites Ausrechnen der Matrix  $P$  ergibt unter Verwendung von (44):

#### Matrix P

$$\begin{aligned} P_{11} &= 1 - \exp(-\lambda\Delta) \\ P_{12} &= 0 \\ P_{13} &= 0 \\ P_{21} &= -(1 - \exp(-\lambda\Delta))\zeta \\ P_{22} &= 1 \\ P_{23} &= -\exp(-\lambda\Delta) \\ P_{31} &= -2 \exp(-\lambda\Delta)(1 - \exp(-\lambda\Delta))\zeta \\ P_{32} &= 0 \\ P_{33} &= 1 - \exp(-2\lambda\Delta) \end{aligned}$$

Im letzten Schritt erhalten wir die Verteilung des Schätzers mit der Matrix T folgendermaßen (vgl. Satz 4.4):

Der Schätzer  $\hat{\gamma}_N = (\kappa_N, \theta_N, \sigma_N)$  ist asymptotisch normal, nämlich mit

$$\sqrt{N}[\hat{\gamma}_N - \gamma_0] \xrightarrow{D} N(0, T)$$

mit  $N \rightarrow \infty$ , wobei

$$T = D\Sigma D^T$$

Hierbei sind die Einträge der Matrix  $D$  gegeben durch (49).

$$D_{ij} = \frac{\partial h_i}{\partial \xi_{j,N}}(\xi_N) \quad i = 1 \dots 3 \quad (49)$$

Die Funktionen  $h_i$  sind Schätzfunktionen der drei Parameter, wobei diese Funktionen nur von  $\xi_N$  abhängen. Dafür werden analog zum vorherigen Modell in den Schätzern der Parameter  $v_{1,N}$  durch  $\xi_{1,N}$  und  $v_{2,N}$  durch  $\xi_{3,N}$  ersetzt.

Wieder führen wir die explizite Berechnung durch mithilfe von (45):

### Matrix D

Die Funktion  $h_1$  schätzt den Parameter  $\lambda$ . Während beim CIR Modell die Wahl der zwei weiteren Schätzfunktionen offensichtlich war, haben wir hier zwei Möglichkeiten:

Entweder wir schätzen  $\nu$  und  $\alpha$ , die Parameter der Gamma Verteilung, oder wir gehen über auf  $\zeta$  und  $\eta$ , welche Erwartungswert und Varianz der Gamma Verteilung widerspiegeln. Offensichtlich können die zwei Parameter auf einfache Weise aus der jeweilig anderen Darstellung gewonnen werden.

Während beide Zugänge im Zuge der Arbeit betrachtet wurden, stellte sich heraus, dass der Zugang mit den Parametern  $\zeta$  und  $\eta$  vorteilhaft ist. Hier ist nämlich der Vergleich mit dem CIR Modell besser gegeben, da der Parameter  $\zeta$  als stationäres Mittel dem dort vorhandenen Parameter  $\theta$  entspricht.

Demzufolge hier also die Schätzfunktionen für die Parameter  $\lambda, \zeta$  und  $\eta$ :

$$\begin{aligned} \hat{\lambda} = h_1(\xi_N) &= -\frac{1}{\Delta} \log \left( \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2} \right), \\ \hat{\zeta} = h_2(\xi_N) &= \frac{\xi_{1,N} - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2} \xi_{1,N}}{1 - \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2}} = \xi_{1,N}, \\ \hat{\eta} = h_3(\xi_N) &= \frac{\xi_{3,N} - \xi_{1,N}^2 - \left( \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2} \right)^2 (\xi_{3,N} - \xi_{1,N}^2)}{1 - \left( \frac{\xi_{2,N} - \xi_{1,N}^2}{\xi_{3,N} - \xi_{1,N}^2} \right)^2} \end{aligned}$$

Damit ergeben sich die gewünschten Ableitungen und somit die Einträge der Matrix  $D$  als:

$$\begin{aligned}
D_{11} &= \frac{\partial}{\partial \xi_{1,N}} h_1(\xi_N) = \frac{1}{\Delta} \frac{2\xi_{1,N}}{\xi_{2,N} - \xi_{1,N}^2} - \frac{1}{\Delta} \frac{2\xi_{1,N}}{xi_{3,N} - \xi_{1,N}^2} \\
D_{12} &= \frac{\partial}{\partial \xi_{2,N}} h_1(\xi_N) = -\frac{1}{\Delta} \frac{\xi_{1,N}}{\xi_{2,N} - \xi_{1,N}^2} \\
D_{13} &= \frac{\partial}{\partial \xi_{3,N}} h_1(\xi_N) = \frac{1}{\Delta} \frac{\xi_{1,N}}{\xi_{3,N} - \xi_{1,N}^2} \\
D_{21} &= \frac{\partial}{\partial \xi_{1,N}} h_2(\xi_N) = 1 \\
D_{22} &= \frac{\partial}{\partial \xi_{2,N}} h_2(\xi_N) = 0 \\
D_{23} &= \frac{\partial}{\partial \xi_{3,N}} h_2(\xi_N) = 0 \\
D_{31} &= \frac{\partial}{\partial \xi_{1,N}} h_3(\xi_N) = -2\xi_{1,N} \\
D_{32} &= \frac{\partial}{\partial \xi_{2,N}} h_3(\xi_N) = 0 \\
D_{33} &= \frac{\partial}{\partial \xi_{3,N}} h_3(\xi_N) = 1
\end{aligned}$$

## 8.2 Bemerkung zur Varianzschätzung Exponentielle Momentenmethode

Aufgrund der Änderung dieser Methode, damit diese funktioniert, ist die bekannte Theorie hierbei nicht anwendbar.

### 8.3 Varianzschätzung Quasi-Maximum-Likelihood Methode

Im Falle der QML Methode ist analog zur ML Methode ein einfacher Schätzer für die Varianzen, vgl. [7]:

$$\left( \frac{\partial}{\partial \gamma} \log(\eta(\bar{\mu}_1, \bar{\sigma}_1^2; \gamma)) \frac{\partial}{\partial \gamma} \log(\eta(\bar{\mu}_1, \bar{\sigma}_1^2; \gamma))^T + \sum_{j=2}^N \frac{\partial}{\partial \gamma} g(r_j, r_{j-1}; \gamma) \frac{\partial}{\partial \gamma} g(r_j, r_{j-1}; \gamma)^T \right)^{-1}$$

wobei  $g(r_j, r_{j-1}; \gamma) := \log(\eta(V_j, \mu(V_{j-1}), \sigma^2(V_{j-1}); \gamma))$  für  $j > 1$  ist.  $\eta$  bezeichnet hier wie bereits bei der Parameterschätzung die Dichte der Normalverteilung.

Zum Bestimmen der Ableitungen verwenden wir die bereits hergeleitete Formel aus (27) und brauchen damit nur noch die jeweiligen Ableitungen nach den Parametern der Funktionen  $\mu$  und  $\sigma^2$  zu bestimmen. Der erste Summand hängt von der stationären Verteilung ab, man muss diesen deshalb auch in Bezug auf die Ableitungen separat betrachten.

Man beachte, dass auch hier als Parameter  $\lambda, \zeta$  und  $\eta$  betrachtet wurden und somit auch dessen Ableitungen angegeben werden:

$$j = 1 : \quad \begin{array}{llll} \mu_1 = \zeta & \frac{\partial}{\partial \lambda} \mu_1 = 0 & \frac{\partial}{\partial \zeta} \mu_1 = 1 & \frac{\partial}{\partial \eta} \mu_1 = 0 \\ \sigma_1^2 = \eta & \frac{\partial}{\partial \lambda} \sigma_1^2 = 0 & \frac{\partial}{\partial \zeta} \sigma_1^2 = 0 & \frac{\partial}{\partial \eta} \sigma_1^2 = 1 \end{array}$$

$j > 1 :$

$$\mu(V_j) = \exp(-\lambda\Delta)v + (1 - \exp(-\lambda\Delta))\zeta$$

$$\frac{\partial}{\partial\lambda}\mu(V_j) = -\exp(-\lambda\Delta)v\Delta + \exp(-\lambda\Delta)\zeta\Delta$$

$$\frac{\partial}{\partial\zeta}\mu(V_j) = 1 - \exp(-\lambda\Delta)$$

$$\frac{\partial}{\partial\eta}\mu(V_j) = 0$$

$$\sigma(V_j) = (1 - \exp(-2\lambda\Delta))\eta$$

$$\frac{\partial}{\partial\lambda}\sigma^2(V_j) = \eta \exp(-2\lambda\Delta)2\Delta$$

$$\frac{\partial}{\partial\zeta}\sigma^2(V_j) = 0$$

$$\frac{\partial}{\partial\eta}\sigma^2(V_j) = 1 - \exp(-2\lambda\Delta)$$

## 9 Vergleich und numerische Illustrationen

### Gamma-OU Modell

Wir vergleichen nun auch bei diesem Modell die drei verschiedenen Methoden zur Parameterschätzung. Einerseits werden die theoretischen Varianzschätzer präsentiert und zudem werden die praktischen Schätzer, die durch wiederholtes Schätzen der Parameter aus mehrfacher Simulation gewonnen wurden, in Form von Histogrammen dargestellt.

Man beachte, dass es bei der exponentiellen Momentenmethode keinen theoretischen Schätzer gibt, aufgrund der Änderung der Methode. Zudem sei noch einmal daran erinnert, dass, im Unterschied zum CIR Modell, die Maximum-Likelihood Methode nicht angewendet werden konnte.

Wir wählen dieselben Parameter wie in der anfänglichen Illustration, wobei wir jetzt jedoch mehr Zeitschritte betrachten. Es sei jedoch hier zu beachten, dass wir zur besseren Vergleichbarkeit mit dem CIR Modell nicht die Parametrisierung  $\gamma = (\lambda, \nu, \alpha)$  betrachten, sondern jene mit  $\gamma = (\lambda, \zeta, \eta)$ , wobei - wie bereits mehrfach erwähnt - diese im folgenden Zusammenhang stehen und leicht auseinander gewonnen werden können:  $\zeta = \frac{\nu}{\alpha}, \eta = \frac{\nu}{\alpha^2}$ . Damit gilt mit der vorherigen Wahl von  $\nu = 53.3334$  und  $\alpha = 8.8889$  also:

$$\lambda = 0.4, \zeta = 6, \eta = 0.6749992, \Delta = 1, N = 10000$$

Die Anzahl der Wiederholungen bei der Simulation und anschließendem Schätzen für die Histogramme ist  $M = 100$ .

#### Momentenmethode

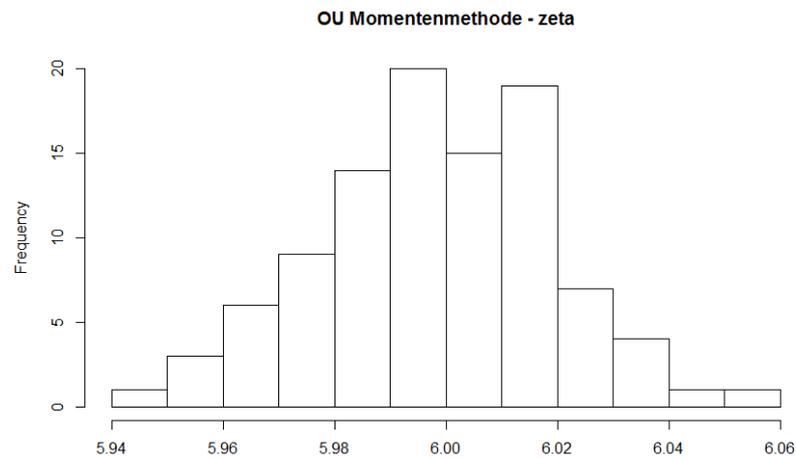
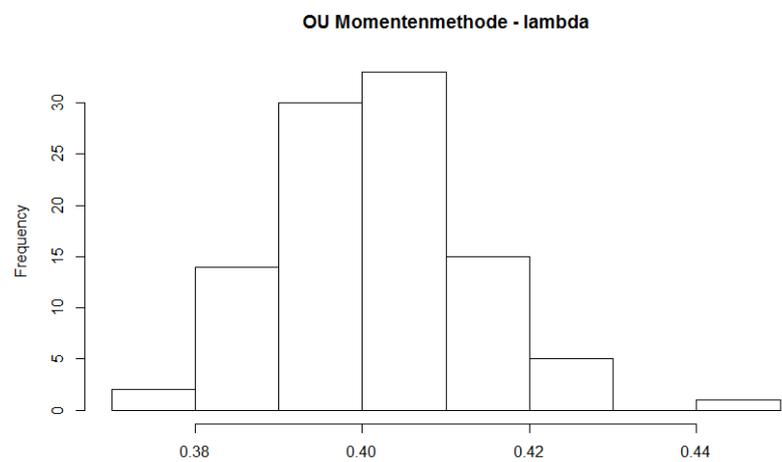
Bei dieser Methode sind die Einträge der Varianz-Kovarianz Matrix recht klein. An den Histogrammen sieht man, dass die Schätzer der Parameter dennoch sehr gut sind. Der Rechenaufwand bei dieser Methode ist zudem gering, was ebenfalls ein Vorteil ist.

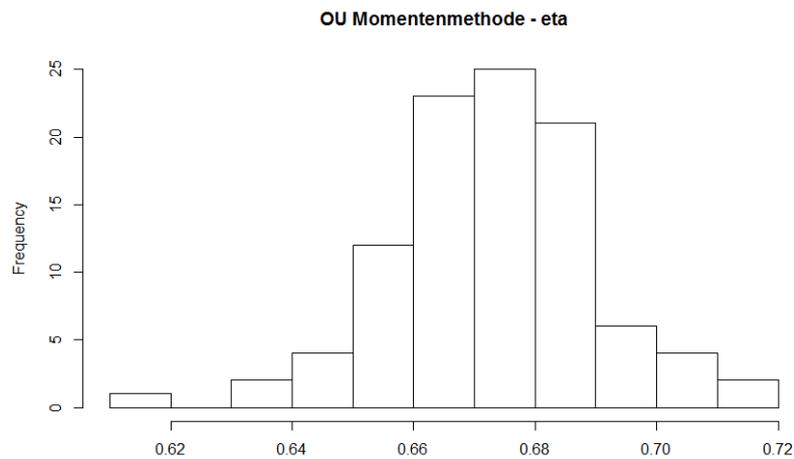
Der theoretische Varianzschätzer, berechnet mit der Methode aus Kapitel 8.1,

ist durch folgende Matrix gegeben:

$$\begin{bmatrix} 1.26234062 & 0.02143533 & -1.3694736 \\ 0.02143533 & 3.41987618 & 0.5532102 \\ -1.36947357 & 0.55321018 & 2.5228064 \end{bmatrix}$$

Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:

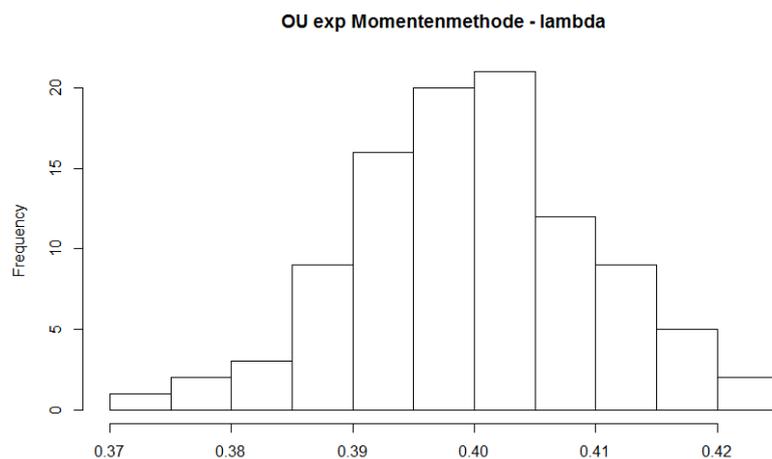


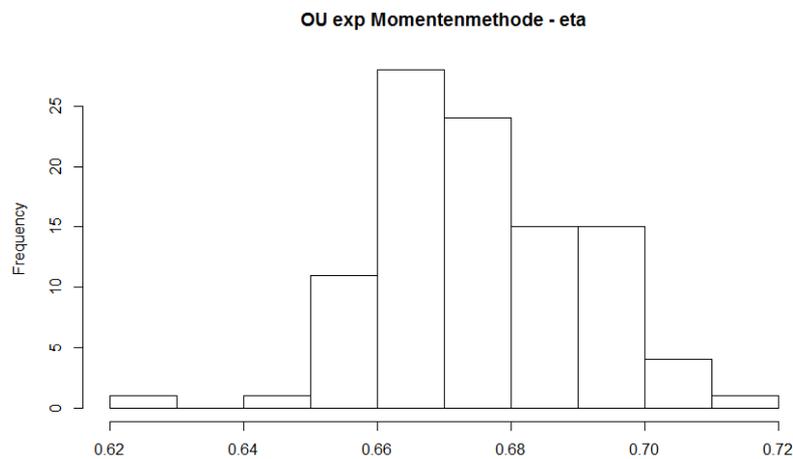
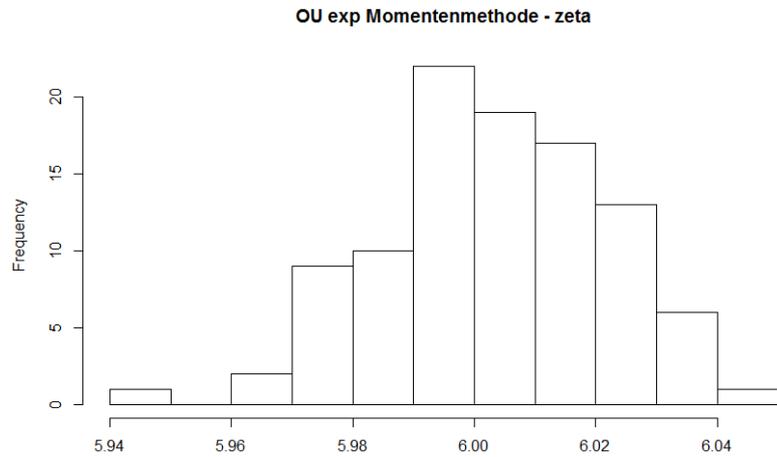


## Exponentielle Momentenmethode

Wie bereits angemerkt, kann hier kein theoretischer Varianzschätzer angegeben werden, weil die Methode geändert wurde. An den Histogrammen sieht man jedoch, dass die Methode ebenfalls sehr gute Schätzer für die Parameter liefert. Die Abweichung von den wahren Parameterwerten ist gering.

Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:





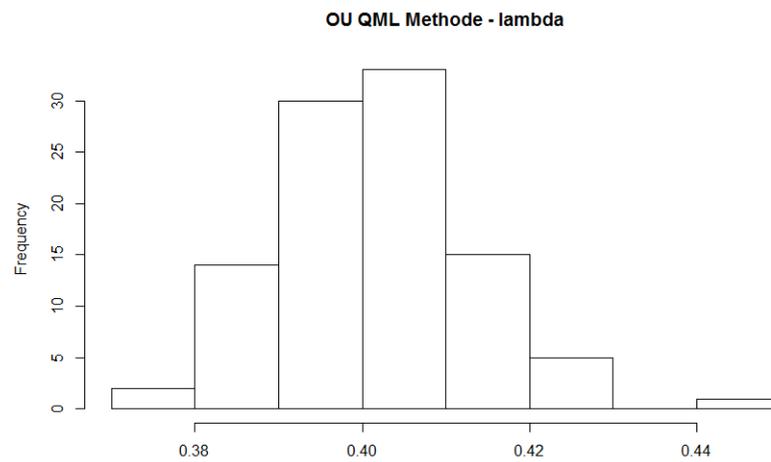
### Quasi-Maximum-Likelihood Methode

Bei dieser Methode hat die Varianz-Kovarianz Matrix kleine Einträge. Ein deutlicher Unterschied zu den anderen Methoden ergibt sich bei Betrachtung der Histogramme, wo man merkt, dass der Parameter  $\eta$  konsequent überschätzt wird. Während der eigentliche Wert  $\eta = 0.6749992$  sein sollte, schwankt der Schätzwert zwischen 1.08 und 1.13. Bei den anderen beiden Parametern ist die Schätzung jedoch sehr gut, wie auch schon bei den vorherigen Methoden.

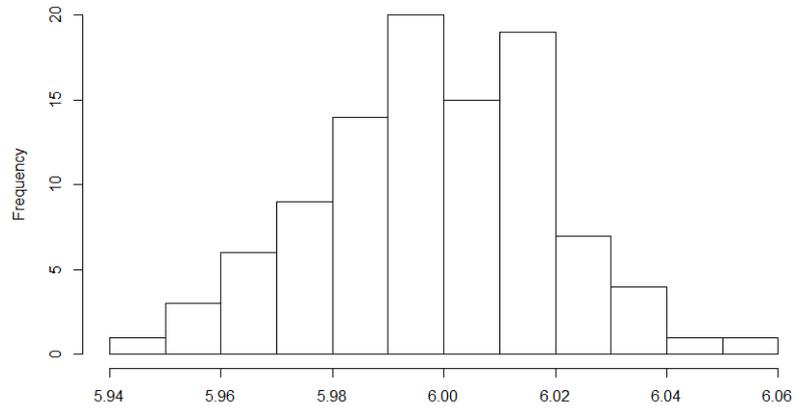
Der theoretische Varianzschätzer, berechnet mit der Methode aus Kapitel 8.3, ist durch folgende Matrix gegeben:

$$\begin{bmatrix} 1.262345282 & -0.003009201 & -1.4097190 \\ -0.003009201 & 3.557452830 & -0.1800659 \\ -1.409718970 & -0.180065864 & 1.8774085 \end{bmatrix}$$

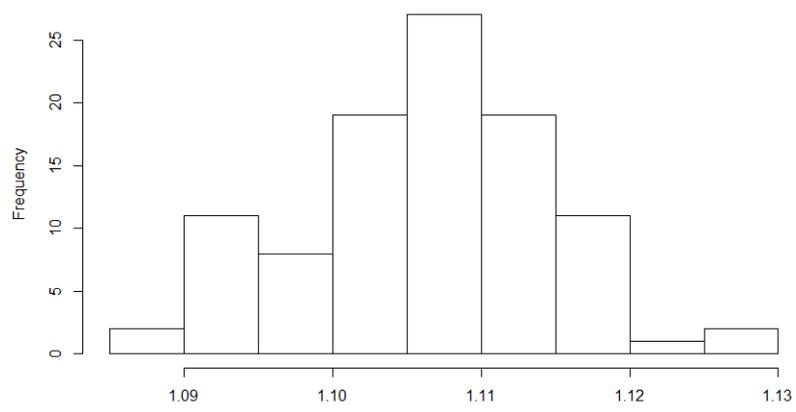
Die Histogramme zu den jeweiligen Parametern sehen folgendermaßen aus:



**OU QML Methode - zeta**



**OU QML Methode - eta**



## Literatur

- [1] Abramowitz, Milton; Stegun, Irene A. *Handbook of mathematical functions with formulas, graphs, and mathematical tables* For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. (1964), 376
- [2] Barndorff-Nielsen, Ole E.; Mikosch, Thomas; Resnick, Sidney I. *Lévy Processes Theory and Applications* Birkhäuser (2001)
- [3] Billingsley, Patrick *Statistical Inference for Markov Processes* The University of Chicago Press, Chicago, Ill. (1961)
- [4] Cont, Rama; Tankov, Peter *Financial Modelling With Jump Processes* Chapman & Hall / CRC financial mathematics series (2004)
- [5] Cox, John C.; Ingersoll JR., Jonathan E.; Ross, Stephen A. *A Theory of the Term Structure of Interest Rates*. *Econometrica*, Volume 53, Issue 2 (Mar. 1985), pages 385 - 408.
- [6] Eberlein, Ernst; Raible, Sebastian *Term Structure Models driven by General Lévy Processes*. *Mathematical Finance*, Vol.9, No. 1 (1999), pages 31-53
- [7] Heyde, Christopher C. *Quasi-likelihood and its application : a general approach to optimal parameter estimation*. New York, NY [u.a.] : Springer (1997)
- [8] Hubalek, Friedrich; Posedel, Petra *Asymptotic Analysis and Explicit Estimation of a Class of Stochastic Volatility Models with Jumps using the martingale Estimating Function Approach*. *Glasnik Matematički*, Vol 48(68) (2013), pages 185 - 210
- [9] Johnson, Norman L. and Kotz, Samuel and Balakrishnan, N. *Continuous univariate distributions. Vol. 2*. Second edition, John Wiley & Sons, Inc., New York (1995), pages 433ff
- [10] Keller-Ressel, Martin *Affine Processes*. Workshop on Interest Rates and Credit Risk, TU Chemnitz, 23. Nov. 2011