

Die geometrische Brownsche Bewegung und Anwendungen

Diese Bachelorarbeit wurde vorgelegt am

Fachbereich 9

Medizintechnik und Technomathematik

FH Aachen, Campus Jülich

von

Fabian Schuller

Matrikelnummer: 3646801

und wurde betreut von

Erstprüfer: Prof. Dr. habil. Daniel Gaigall

Zweitprüfer: Thorsten Adrian, MSc

6. Oktober 2025

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die Bachelorarbeit mit dem Titel

Die geometrische Brownsche Bewegung und Anwendungen

selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, alle Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, kenntlich gemacht sind und die Arbeit in gleicher oder ähnlicher Fassung noch nicht Bestandteil einer Studien- oder Prüfungsleistung war. Ich verpflichte mich, ein Exemplar der Bachelorarbeit fünf Jahre aufzubewahren und auf Verlangen dem Prüfungsamt des Fachbereiches Medizintechnik und Technomathematik auszuhändigen.

Aachen, den 6. Oktober 2025



Fabian Schuller

Inhaltsverzeichnis

1	Motivation	4
2	Notation	4
3	Stochastische Prozesse	5
3.1	Bedingter Erwartungswert und Filtrationen	6
3.2	Eigenschaften von stochastischen Prozessen	9
4	Die Brownsche Bewegung	10
4.1	Diskrete Brownsche Bewegung	10
4.2	Nützliche Ergebnisse aus der Stochastik	13
4.3	Die Brownschen Bewegung als Grenzprozess und Eigenschaften .	14
5	Die geometrische Brownsche Bewegung	20
5.1	Die geometrische Brownschen Bewegung	20
5.2	Die logarithmische Normalverteilung	22
6	Anwendungen auf Zeitreihen	22
6.1	Kalibrierung	22
6.2	Bootstrap-Verfahren zur Kalibrierung	23
6.3	Berechnung von Konfidenzintervallen	24
6.4	Back-Tests	25
6.5	Monte-Carlo-Simulation	28
7	Aktienoptionen	29
7.1	Finanzmathematische Grundlagen	29
7.2	Bewertung von Aktienoptionen im Binomialmodell mit dem Risikoneutralen Maß	30
7.3	Das Black-Scholes-Modell als Grenzfall des Binomialmodells . . .	34
7.4	Bewertung von Aktienoptionen mit Monte-Carlo-Verfahren . . .	39
8	Alternative Modelle	42
8.1	Stochastische Differentialgleichungen	43
8.2	Charakterisierung alternativer Kursmodelle durch stochastische Differentialgleichungen	45
8.3	Simulation stochastischer Differentialgleichungen	46
8.4	Implementierung eines CEV (Constant Elasticity of Variance) Modell	47
8.5	Ergebnisse und Vergleich der Modelle	52
9	Fazit	55
10	Quellenverzeichnis	56

1 Motivation

Relevanz Die Brownsche Bewegung ist ein zentrales Konzept in der Stochastik und findet Anwendung in den Naturwissenschaften, und insbesondere in der Finanzmathematik. Eine Modifikation ist die geometrische Brownsche Bewegung. Sie dient als Grundlage für die Modellierung von Aktienkursen und anderen finanziellen Zeitreihen. Eine Vielzahl von aktuellen Aktienkurs-Modellen basieren auf dieser Theorie.

Ablauf Die Arbeit beginnt mit grundlegender Notation und Begriffen (Kapitel 2) und führt anschließend in stochastische Prozesse ein (Kapitel 3), inklusive Zufallsspaziergang, Binomialmodell sowie Filtration, bedingtem Erwartungswert, Markov- und Martingaleigenschaften. Kapitel 4 konstruiert die Brownsche Bewegung als Grenzprozess diskreter Normalverteilungssummen und diskutiert zentrale Eigenschaften wie Stetigkeit und Selbstähnlichkeit. Darauf aufbauend wird in Kapitel 5 die geometrische Brownsche Bewegung aus dem Binomialmodell via Taylor-Approximation und Grenzwertsätzen hergeleitet und die Lognormalverteilung von Kursen gezeigt. Kapitel 6 widmet sich Anwendungen auf Zeitreihen: Kalibrierung von μ und σ , Konfidenzintervalle/-bänder, Backtests und Fehlermaße sowie Monte-Carlo-Simulation. In Kapitel 7 werden Aktienoptionen behandelt: finanzmathematische Grundlagen, Bewertung im (erweiterten) Binomialmodell unter risikoneutralem Maß, der Grenzfall zur Black-Scholes-Formel sowie Monte-Carlo-Bewertung (auch pfadabhängig). Kapitel 8 schließt mit Fazit, methodischer Einordnung, weiterführenden Aspekten und Ausblick.

2 Notation

- $E(X)$ bzw. $E[X]$ - Erwartungswert
- $V(X)$ bzw. $V[X]$ - Varianz
- $X \sim N(\mu, \sigma^2)$ - normalverteilte Zufallsvariable mit Erwartungswert μ und Varianz σ^2
- (Ω, \mathcal{F}, P) - Wahrscheinlichkeitsraum
- W_t - Brownsche Bewegung (Wiener-Prozess)
- GBM - geometrische Brownsche Bewegung
- CEV - Constant Elasticity of Variance (Modell)
- MSE - Mean Squared Error
- NRMSE - Normalized Root Mean Squared Error
- RMSE - Root Mean Squared Error
- MAPE - Mean Absolute Percentage Error

Konvergenzbegriffe

- $X_n \rightarrow X$ - punktweise Konvergenz (Für jedes $\omega \in \Omega$ gilt: $X_n(\omega) \rightarrow X(\omega)$)
- $X_n \xrightarrow{\text{glm.}} X$ - gleichmäßige Konvergenz (Für jede $\varepsilon > 0$ gilt: $\sup_{\omega \in \Omega} |X_n(\omega) - X(\omega)| < \varepsilon$ für n groß genug)
- $X_n \xrightarrow{d} X$ - Konvergenz in Verteilung (Die Verteilungsfunktion F_{X_n} von X_n konvergiert punktweise gegen die Verteilungsfunktion F_X von X , mindestens an den Stetigkeitsstellen von F_X)
- $X_n \xrightarrow{p} X$ - Konvergenz in Wahrscheinlichkeit (Für jede $\varepsilon > 0$ gilt: $P(|X_n - X| > \varepsilon) \rightarrow 0$)
- $X_n \xrightarrow{\text{f.s.}} X$ - fast sichere Konvergenz (Für fast alle $\omega \in \Omega$ gilt: $X_n(\omega) \rightarrow X(\omega)$)

Auf das Verhältnis zwischen den Konvergenzbegriffen wird bei Bedarf eingegangen.

3 Stochastische Prozesse

In einer einführenden Vorlesung zur Stochastik betrachtet man zunächst einzelne Zufallsvariablen, etwa das Ergebnis eines Würfelwurfs. In dieser Arbeit stehen hingegen Folgen von Zufallsvariablen im Mittelpunkt. Solche Folgen erlauben es, zeitliche Entwicklungen zu modellieren, beispielsweise Veränderungen eines Systems, die zu bestimmten Zeitpunkten t_0, t_1, \dots gemessen werden. Man spricht dann von einem zeitdiskreten stochastischen Prozess. Eine natürliche Verallgemeinerung bilden zeitstetige stochastische Prozesse: Hierbei definiert man eine Familie von Zufallsvariablen $(X_t)_{t \in \mathbb{R}_{\geq 0}}$, wobei t kontinuierlich als Zeitparameter interpretiert wird. Ziel dieser Arbeit ist es, Aktienkurse und später Optionswerte mit einem stochastischen Prozess zu modellieren.

Beispiel 3.0.1 (Zufallsspaziergang). Ein einfacher stochastischer Prozess ist der *Zufallsspaziergang* (random walk). Die Position des Spaziergängers zur Zeit t wird durch eine Zufallsvariable X_t beschrieben. Man startet bei $X_0 = 0$. In jedem Zeitschritt bewegt sich der Spaziergänger entweder einen Schritt nach rechts oder nach links, jeweils mit Wahrscheinlichkeit p bzw. $1 - p$. Formal gilt:

$$X_{t+1} = X_t + \xi_{t+1},$$

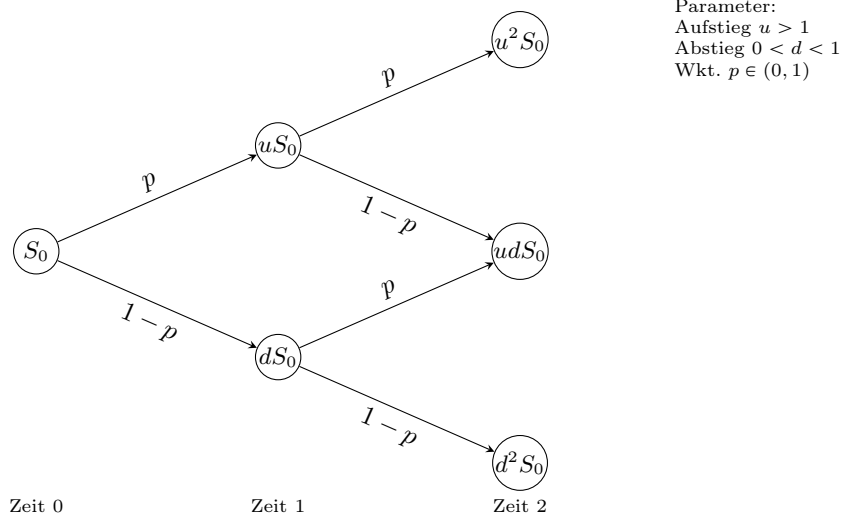
wobei ξ_{t+1} eine unabhängige Zufallsvariable ist mit

$$\xi_{t+1} = \begin{cases} +1, & \text{mit Wahrscheinlichkeit } p, \\ -1, & \text{mit Wahrscheinlichkeit } 1 - p. \end{cases}$$

Beispiel 3.0.2 (Binomialmodell). Das Binomialmodell ist ein simples Modell einer Aktie und deren Preisentwicklung. Man beginnt mit einem Anfangspreis S_0 . S_1, S_2, \dots sind dann Messungen des Aktienpreises zu einem festen Intervall. Zu einer festen Wahrscheinlichkeit $0 < p < 1$ steigt die Aktie um den Faktor u , oder fällt mit der Wahrscheinlichkeit $1 - p$ um den Faktor d . Also

$$S_{t+1} = \begin{cases} d \cdot S_t, & \text{mit Wahrscheinlichkeit } p \\ u \cdot S_t, & \text{mit Wahrscheinlichkeit } 1 - p \end{cases}.$$

Im Gegensatz zum Zufallsspaziergang sind die Schritte hier nicht additiv, sondern multiplikativ. Daher eignet sich das Binomialmodell (im Gegensatz zum Zufallsspaziergang) zur Modellierung von Aktienkursen, es sind keine negativen Preise möglich.



Beispiel 3.0.3 (Brownsche Bewegung). Die Brownsche Bewegung beschreibt die Bewegung eines Partikels in einer Flüssigkeit [5]. Die Zufallsgröße ist hier die Position des Partikels. Naturanaloge stochastische Prozesse erfüllen oft eine Stetigkeitsbedingung: Andernfalls würde sich z. B. der Partikel in der Flüssigkeit teleportieren. Konkret: Es bezeichne der Pfad $t \mapsto X_t$ die Realisierungen der Zufallsvariablen X_t unter der Zeit t . Dann wird gefordert, dass die Pfade fast-sicher stetig sind, also

$$P(\{\omega \in \Omega | t \mapsto X_t(\omega) \text{ ist stetig}\}) = 1.$$

Im nächsten Kapitel wird die Brownsche Bewegung formal definiert.

3.1 Bedingter Erwartungswert und Filtrationen

Im Folgenden wird stochastische Unabhängigkeit und der Erwartungswert im Bezug auf die Zeit untersucht. Dazu definiert man einen neuen Wahrscheinlichkeitsraum, der alle möglichen Verläufe des Prozesses vereint. Um nicht weiter

zwischen zeit-stetigen und -diskreten Prozessen unterscheiden zu müssen, sei $I = \mathbb{N}_0$ im diskreten und $I = \mathbb{R}_{\geq 0}$ im stetigen Fall. Des weiteren sei $(\tilde{\Omega}, \mathcal{A}_t, \tilde{P})$ ein Wahrscheinlichkeitsraum, und $(X_t)_{t \in I}$ eine Familie von Zufallsvariablen auf dem Wahrscheinlichkeitsraum, die einen stochastischen Prozess bilden. Vor-erst wird der Prozess auf dem Produkt-Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{F}, P) := \bigtimes_{t \in I} (\tilde{\Omega}, \mathcal{A}_t, \tilde{P})$$

betrachtet.

Definition 3.1.1 (Adaptiertheit). In der modernen Wahrscheinlichkeitsrechnung wird Information über den Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) als Teil- σ -Algebra von \mathcal{A} verschlüsselt (Behrendts [2] S. 13). Dazu wird der Begriff der Filtration definiert: Eine Familie von σ -Algebren $\mathcal{F}_t, t \in I$ heißt *Filtration*, wenn $\mathcal{F}_s \subset \mathcal{F}_t$ für alle $s < t$ gilt. Der Prozess $(X_t)_{t \in I}$ heißt an die Filtration adaptiert, wenn X_t \mathcal{F}_t -messbar ist für alle $t \in I$. Die Eigenschaft $\mathcal{F}_s \subset \mathcal{F}_t$ kann man wie folgt interpretieren: Die Größe des Datensatzes nimmt über die Zeit zu. Anstatt dem Produktraum, kann man die Zufallsvariablen X_t somit auf den Räumen $(\Omega, \mathcal{F}_t, P)$ definieren.

Beispiel 3.1.2 (Filtration des Binomialmodells). Jeder Zeitschritt im Binomialmodell ist zunächst eine Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\{W, B\}, \sigma(\{W, B\}), P)$ mit $P(\{W\}) = p$, $P(\{B\}) = 1 - p$. Hier steht W für den Kursanstieg und B für Kursfall. Die Preisentwicklung bis zum Zeitpunkt t_0 fasst man als Produkt-Wahrscheinlichkeitsraum auf. Betrachtet man ein dreistufiges Binomialmodell, ergibt sich die folgende Filtration:

$$\begin{aligned} \mathcal{F}_0 &= \{\emptyset, \Omega\} \\ \mathcal{F}_1 &= \{\{WWW, WWB, \dots, WBB\}, \{BWW, BWB, \dots, BBB\}, \emptyset, \Omega\} \\ \mathcal{F}_2 &= \{\{WWB, WWW\}, \{WBB, WBW\}, \{BWB, BWW\}, \{BBB, BBW\}, \\ &\quad \{WWW, WWB, \dots, WBB\}, \{BWW, BWB, \dots, BBB\}, \\ &\quad \emptyset, \Omega\} \\ \mathcal{F}_3 &= \text{Pot}(\Omega) \end{aligned}$$

Der Prozess S_t ist adaptiert: Im ersten Schritt kann der Kurs entweder steigen oder fallen, alle weiteren Kursverläufe sind in dem Ereignis enthalten.

Definition 3.1.3 (Wahrscheinlichkeitsraum eines stochastischen Prozesses). Mit der Filtration kann man den Wahrscheinlichkeitsraum eines stochastischen Prozesses genauer beschreiben. Der Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , wobei $\mathcal{F} = \lim_{t \rightarrow \infty} \mathcal{F}_t$, entsteht durch die Filtration \mathcal{F}_t . Der stochastische Prozess existiert dann im filtrierte Wahrscheinlichkeitsraum. Man schreibt

$$(\Omega, \mathcal{F}_t, \mathcal{F}, P).$$

Definition 3.1.4 (Bedingter Erwartungswert). Der bedingte Erwartungswert ist wichtig für die Untersuchung stochastischer Prozesse, insbesondere für

unseren Anwendungsfall, da historische Entwicklungen in der Praxis meist bekannt sind. Beispiel Poker: Dem Spieler ist seine Hand, und die Karten auf dem Tisch bekannt. Daraus lässt sich die eigene Gewinnwahrscheinlichkeit berechnen. Die Chips auf dem Tisch bilden eine untere Grenze für den erwarteten Gewinn:

$$E(\text{Chips dieser Runde} | \text{bekannte Karten}) \geq P(\text{Gewinn} | \text{bekannte Karten}) \cdot \text{Chips auf dem Tisch}$$

Sei B ein Ereignis mit $P(B) \neq 0$. Dann definiert man den *bedingten Erwartungswert* als

$$E(X_t | B) := \frac{1}{P(B)} \int_B X_t(\omega) dP(\omega),$$

und für einen diskreten Wahrscheinlichkeitsraum reicht

$$E(X_t | B) := \frac{1}{P(B)} \sum_{b \in B} X_t(b) \cdot P(\{b\})$$

Mit dem bedingten Erwartungswert wird die Zufallsvariable $E(X_t)$. Für eine Zufallsvariable Y wird durch den bedingten Erwartungswert $E(X_t | Y)$ eine neue Zufallsvariable definiert:

$$E(X_t | Y)(\omega) := E(X_t | Y = Y(\omega)).$$

Im Folgenden werden die Sätze des totalen Erwartungswertes und des iterierten Erwartungswertprinzips bewiesen, allerdings nur im diskreten Fall. Das ist ausreichend, da die meisten stetigen Prozesse als Grenzprozesse diskreter Prozesse aufgefasst werden können. Dann werden die folgenden Aussagen ggf. durch Grenzwertsätze auf den stetigen Fall übertragen.

Beispiel 3.1.5 (Rechenbeispiel). Angenommen, die Kursentwicklung im dreistufigen Binomialmodell ist bis zum Zeitpunkt $t = 1$ bekannt, nämlich ist der Preis um den Faktor d gestiegen. Was kann man im Schritt $t = 2$ erwarten? Hier ist $B = \{WWW, WWB, \dots, WBB\}$, $P(B) = p$, und gesucht ist $E(S_2 | B)$. Aus der Definition folgt

$$\begin{aligned} E(S_2 | B) &= \frac{1}{P(B)} \sum_{k \in B} S_2(k) \cdot P(\{k\}) \\ &= \frac{1}{p} (S_2(WWW) \cdot P(WWW) + \dots + S_2(WBB) \cdot P(WBB)) \\ &= \frac{1}{p} (S_0 \cdot d^2 \cdot p^2 + \dots + S_0 \cdot du \cdot p(1-p)) \end{aligned}$$

Für den Startwert $S_0 = 10$ und $p = 0.25$ so wie $d = \frac{1}{u} = 2$ ergibt sich

$$E(S_2 | B) = 4 \cdot 10 \cdot (2^2 \cdot 0.25^2 + 2^2 \cdot 0.25^2 + 1 \cdot 0.25 \cdot 0.75 + 1 \cdot 0.25 \cdot 0.75) = 35$$

Ist die bisherige Kursentwicklung bekannt, hier $S_0 = 10, S_1 = d \cdot S_0 = 20$, aber nicht B , müsste man zuerst $B := S_1^{-1}(20)$ berechnen.

Satz 3.1.6 (Totaler Erwartungswert). Der Satz vom totalen Erwartungswert besagt, dass der Erwartungswert einer Zufallsvariablen durch die Summe der bedingten Erwartungswerte über eine Partition des Wahrscheinlichkeitsraumes berechnet werden kann. Sei also $(A_i)_{i \in I}$ eine Partition von Ω (also A_i paarweise disjunkt, $\bigcup_{i \in I} A_i = \Omega$), dann gilt

$$E(X_t) = \sum_{i \in I} E(X_t | A_i) \cdot P(A_i).$$

In dieser Arbeit wird nur der diskrete Fall betrachtet. *Beweis.*

$$E(X_t) = \sum_{\omega \in \Omega} X_t(\omega) \cdot P(\{\omega\}) = \sum_{i \in I} \sum_{\omega \in A_i} X_t(\omega) \cdot P(\{\omega\}) = \sum_{i \in I} E(X_t | A_i) \cdot P(A_i).$$

□

Satz 3.1.7 (Iterierter Erwartungswert / Turmeigenschaft). Für diskrete Zufallsvariablen X und Y gilt

$$E(E(X | Y)) = E(X).$$

(Wie oben ist $E(X | Y)$ eine Zufallsvariable.) *Beweis.* Nach Definition von $E(X | Y)$ ist

$$E(X | Y) = g(Y) = \sum_x x P(X = x | Y = y).$$

Damit gilt

$$\begin{aligned} E(E(X | Y)) &= \sum_y g(y) P(Y = y) = \sum_y \left(\sum_x x P(X = x | Y = y) \right) P(Y = y) \\ &= \sum_{x,y} x P(X = x | Y = y) P(Y = y) = \sum_{x,y} x P(X = x, Y = y) \\ &= \sum_x x P(X = x) = E(X). \end{aligned}$$

Hier wurde im dritten Schritt $P(X = x | Y = y)P(Y = y) = P(X = x, Y = y)$ benutzt und im letzten Schritt die Summation über y zusammengefasst. □

3.2 Eigenschaften von stochastischen Prozessen

Definition 3.2.1 (Martingal). Martingale sind Prozesse, die tendenziell weder steigen noch fallen, also faire Prozesse. Tendenzial heißt auf den Erwartungswert bezogen. Steigende Prozesse werden Supermartingale genannt, fallende Submartingale. Die mathematische Definition erfolgt durch den Bedingten Erwartungswert: Ist ein fairer Kurs zum aktuellen Zeitpunkt s auf einem bestimmten Wert, liegt der Erwartungswert zur Messzeit $t > s$ bei dem selben Wert.

Ein stochastischer Prozess $(X_t)_{t \in I}$ heißt *Submartingal*, wenn $E(X_t | X_s = v) \leq v$ für alle $s < t$, und alle $v \in \mathbb{R}$. $(X_t)_{t \in I}$ heißt *Supermartingal*, wenn $(-X_t)_{t \in I}$ ein Submartingal ist, und *Matringal*, wenn er sowohl Supermartingal als auch Submartingal ist, also $E(X_t | X_s = v) = v$.

Beispiel 3.2.2 (Martingaleigenschaft des Binomialmodells). Das Binomialmodell genau dann ein Martingal, wenn $p = \frac{1-d}{u-d}$ gilt, wobei $u \neq d$. Ansonsten ist das Binomialmodell ebenfalls ein Martingal falls $p = 0.5$.

Beweis für $u \neq d$. Zuerst wird der Fall $t = s+1$ gezeigt: Es gilt $S_{s+1} = S_s \xi_{s+1}$ für eine Zufallsvariable ξ_{s+1} mit $\mathbb{P}(\xi_{s+1} = u) = p$, $\mathbb{P}(\xi_{s+1} = d) = 1 - p$. Es folgt

$$E(S_{s+1}|S_s = v) = v \cdot E(\xi_{s+1}) = v(pu + (1-p)d).$$

Und setzt man die Martingaleigenschaft $E(S_{s+1}|S_s = v) = v$ ein, ergibt sich

$$pu + (1-p)d = 1 \iff p = \frac{1-d}{u-d}.$$

Seien nun $s < t \in \mathbb{N}_0$ beliebig. Dann gilt $S_t = \xi_{t-1} \cdot \xi_{t-2} \cdots \xi_{s+1} \cdot S_s$. Da die ξ_i unabhängig und identisch verteilt sind, gilt

$$E(S_t|S_s = v) = v \cdot \prod_{i=s+1}^{t-1} E(\xi_i) = v \cdot E(\xi_{s+1})^{t-s} = v \cdot (pu + (1-p)d)^{t-s}.$$

$E(S_t|S_s = v) = v$ ist wieder äquivalent zu $p = \frac{1-d}{u-d}$. □

Definition 3.2.3 (Markovprozess). Ein stochastischer Prozess heißt Markovprozess, falls die Zufallsvariablen X_t lediglich vom unmittelbaren Vorgänger X_s , $s < t$ abhängen. Konkret: Für alle $r < s < t$ und alle $u, v, w \in \mathbb{R}$ gilt

$$P(X_t \leq w | X_s = v, X_r = u) = P(X_t \leq w | X_s = v).$$

Beispiel 3.2.4. Das Binomialmodell und der Zufallsspaziergang sind Beispiele für Markovprozesse, da die zukünftige Position nur von der aktuellen Position und nicht von der gesamten Vergangenheit abhängt.

4 Die Brownsche Bewegung

Oft wird die Brownsche Bewegung (oder auch Wiener Prozess) axiomatisch definiert. In dieser Arbeit werden direkt kumulative Summen von Normalverteilungen betrachtet. Zuerst wird eine vereinfachte Darstellung des Prozesses eingeführt, die diskrete Brownsche Bewegung. Die Argumentation ist angelehnt an Behrends [2] (2013, Kap. 5).

4.1 Diskrete Brownsche Bewegung

Definition 4.1.1 (Diskrete Brownsche Bewegung). Die elementare Brownsche Bewegung sei ein stochastischer Prozess, der aus einer Folge von Zufallsvariablen $\xi_n, n \in \mathbb{N}_0$ besteht, wobei

$$\xi_n = \sum_{i=1}^n \eta_i, \quad \eta_i \sim N(0, 1)$$

. Die Zufallsvariablen η_i sind unabhängig und identisch verteilt. Nun wird eine stetige Zeitentwicklung durch lineare Interpolation eingeführt:

$$b^{(1)}(t) := \xi_{[t]} + (t - [t])(\xi_{[t]+1} - \xi_{[t]}), \quad t \geq 0.$$

Die Funktion $b^{(1)}(t)$ wird diskrete Brownsche Bewegung erster Ordnung genannt. Der Name rührt daher, dass in eine Zeiteinheit genau eine Normalverteilung einbezogen wird. Im Allgemeinen wird die diskrete Brownsche Bewegung $b^{(N)}(t)$ N -ter Ordnung definiert als

$$b^{(N)}(t) := \frac{1}{\sqrt{N}} \left(\xi_{[Nt]} + (Nt - [Nt])(\xi_{[Nt]+1} - \xi_{[Nt]}) \right), \quad t \geq 0.$$

Hierbei wird in eine Zeiteinheit N Normalverteilungen einbezogen. Der Faktor $1/\sqrt{N}$ dient dazu, die Varianzen der Normalverteilungen zu normieren.

Beispiel 4.1.2 (Visualisierung der diskreten Brownschen Bewegung).

Das folgende R-Programm (Ausschnitt) generiert eine diskrete Brownsche Bewegung N -ter Ordnung.

```

1  n_points <- N * T_max
2  eta <- rnorm(n_points, mean=0, sd=1)
3  xi <- c(0, cumsum(eta))
4  t_grid <- seq(0, T_max, length.out=steps)
5  k <- floor(N * t_grid)
6  frac <- N * t_grid - k
7  vals <- xi[k+1] + frac * (xi[k+2] - xi[k+1])
8  vals <- vals / sqrt(N)
```

Für verschiedene Werte von N ergeben sich die folgenden Grafiken:

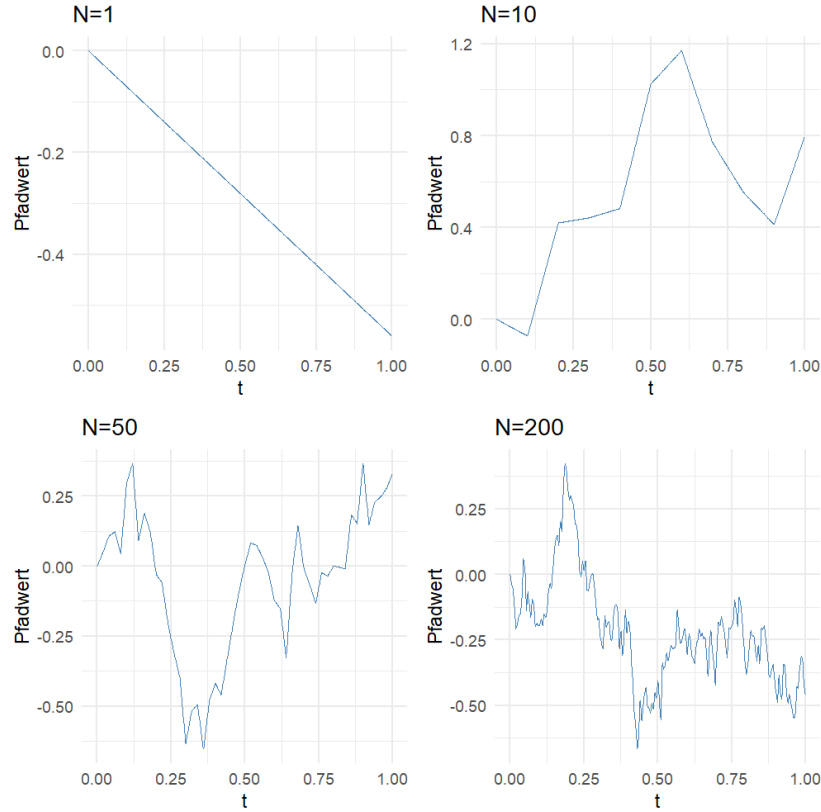


Abbildung 1: Diskrete Brownsche Bewegung erster, zehnter, fünfzigster und zweihundertster Ordnung

Lemma 4.1.3 (Martingal-Eigenschaft der diskreten Brownschen Bewegung). Ohne Beschränkung der Allgemeinheit wird der Fall $N = 1$ betrachtet, sonst kann die Zeit skaliert werden. Ebenso wird nur der Fall $t \in \mathbb{N}_0$ untersucht: Der Prozess ist also wieder diskret. Das ist ausreichend, da im Verlauf der Arbeit der Grenzprozess $N \rightarrow \infty$ relevant wird. Dann ist jeder Wert der Brownschen Bewegung beliebig nah an einem der diskreten Werte, und die Martingal-Eigenschaft folgt aus der Stetigkeit des Erwartungswertes. Der stetig interpolierte diskrete Prozess an sich ist nämlich kein Martingal. *Beweis.* Zuerst wird der Fall $m = n + 1$ betrachtet: zu zeigen ist

$$E(b^{(1)}(m)|b^{(1)}(n) = v) = v.$$

Da $b^{(1)}(m) = b^{(1)}(n) + \eta_{n+1}$ folgt

$$E(b^{(1)}(m)|b^{(1)}(n) = v) = E(v + \eta_{n+1}|b^{(1)}(n) = v) = v + E(\eta_{n+1}) = v.$$

Nun der allgemeine Fall $m > n+1$: aus dem Satz des iterierten Erwartungswertes

folgt

$$E(b^{(1)}(m)|b^{(1)}(n) = v) = E(E(b^{(1)}(m)|b^{(1)}(m-1))|b^{(1)}(n) = v).$$

Da $E(b^{(1)}(m)|b^{(1)}(m-1)) = b^{(1)}(m-1)$ folgt

$$E(b^{(1)}(m)|b^{(1)}(n) = v) = E(b^{(1)}(m-1)|b^{(1)}(n) = v).$$

Induktiv folgt die Behauptung. \square

Lemma 4.1.4 (Varianz der diskreten Brownschen Bewegung).

$$\begin{aligned} V(b^{(N)}(t)) &= \frac{1}{N} (V(\xi_{\lfloor Nt \rfloor}) + (Nt - \lfloor Nt \rfloor)^2 \cdot V(\xi_{\lfloor Nt \rfloor + 1} - \xi_{\lfloor Nt \rfloor})) \\ &= \frac{1}{N} (\lfloor Nt \rfloor + (Nt - \lfloor Nt \rfloor)^2) \\ &= \frac{\lfloor Nt \rfloor}{N} + \frac{(Nt - \lfloor Nt \rfloor)^2}{N}. \end{aligned}$$

Im Grenzübergang $N \rightarrow \infty$ konvergiert $\frac{\lfloor Nt \rfloor}{N} \rightarrow t$ und $\frac{(Nt - \lfloor Nt \rfloor)^2}{N} \rightarrow 0$. \square

4.2 Nützliche Ergebnisse aus der Stochastik

Im Folgenden Beweis zur Stetigkeit der Pfade werden einige Ergebnisse aus der Stochastik verwendet, die hier zusammengefasst, aber nicht bewiesen werden.

Lemma 4.2.1 (Reihenkriterium für fast sichere Konvergenz, Henze [9] S. 201). Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen. Wenn es eine Reihe $\sum_{n=1}^{\infty} a_n < \infty$ mit $a_n \geq 0$ gibt, so dass

$$P(|X_n| > \varepsilon) \leq a_n \quad \text{für alle } n \in \mathbb{N} \text{ und jedes } \varepsilon > 0,$$

dann konvergiert X_n fast sicher gegen 0, d.h.

$$P\left(\lim_{n \rightarrow \infty} X_n = 0\right) = 1.$$

Wird hier nicht bewiesen. \square

Satz 4.2.2 (Cramér-Wold-Technik, Henze [9] S. 225). Seien $(X_n)_{n \in \mathbb{N}}$ und X Zufallsvektoren in \mathbb{R}^d . Dann sind folgende Aussagen äquivalent:

1. $X_n \xrightarrow{d} X$ für $n \rightarrow \infty$.
2. Für alle $t \in \mathbb{R}^d$ gilt $t^T X_n \xrightarrow{d} t^T X$ für $n \rightarrow \infty$.

Wird hier nicht bewiesen. \square

Die Cramér-Wold-Technik erlaubt es, mehrdimensionale Verteilungskonvergenz auf eindimensionale Berechnungen zurückzuführen.

4.3 Die Brownschen Bewegung als Grenzprozess und Eigenschaften

Nun wird der Grenzprozess $N \rightarrow \infty$ betrachtet. Intuitiv wird die Zeit immer feiner aufgelöst, und es werden immer mehr Normalverteilungen in eine Zeiteinheit einbezogen. Vorerst ist jedoch unklar, ob der Grenzprozess überhaupt existiert. Um die Konvergenz des Prozesses zu zeigen, wird die Verteilungskonvergenz untersucht. Konvergenz der einzelnen Zeitpunkte (oder endlich-dimensionalen Vektoren) reicht nicht für Konvergenz der Prozesse als Funktionen. Daher wird die Konvergenz in 3 Schritten untersucht:

1. Verteilungskonvergenz der einzelnen Zeitpunkte
2. Verteilungskonvergenz von endlich-dimensionalen Vektoren und die Kovarianzen (Hieraus folgt bereits die Selbstähnlichkeit / Bedingte Verteilung so wie die Martingal-Eigenschaft)
3. Stetigkeit der Pfade des Grenzprozesses

Lemma 4.3.1 (Verteilungskonvergenz der diskreten Brownschen Bewegung). Es existiert ein stochastischer Prozess $W_t, t \geq 0$, so dass für jedes t die Verteilung von $b^{(N)}(t)$ gegen die Verteilung von W_t konvergiert, wenn $N \rightarrow \infty$. *Beweis.* Da

$$b^{(N)}(t) = \frac{1}{\sqrt{N}}(\xi_k + \alpha(\xi_{k+1} - \xi_k)) = \frac{1}{\sqrt{N}}(\xi_k + \alpha \eta_{k+1}).$$

ist $b^{(N)}(t)$ eine lineare Kombination von Normalverteilungen, und daher wiederum normalverteilt. Erwartungswert und Varianz wurden bereits berechnet:

$$E(b^{(N)}(t)) = 0, \quad V(b^{(N)}(t)) = \sigma_N^2(t) = \frac{1}{N}(k + \alpha^2),$$

wobei $k = \lfloor Nt \rfloor$ und $\alpha = Nt - k \in [0, 1)$. Damit gilt

$$b^{(N)}(t) \sim N\left(0, \frac{1}{N}(k + \alpha^2)\right).$$

Die Verteilungsfunktion ist gegeben durch

$$F_N(x) = P(b^{(N)}(t) \leq x) = \Phi\left(\frac{x}{\sigma_N(t)}\right),$$

wobei Φ die Standardnormalverteilungsfunktion ist. Aus $\sigma_N^2(t) \rightarrow t$ folgt $\sigma_N(t) \rightarrow \sqrt{t}$ und wegen der Stetigkeit von Φ daher

$$F_N(x) = \Phi\left(\frac{x}{\sigma_N(t)}\right) \longrightarrow \Phi\left(\frac{x}{\sqrt{t}}\right) \quad \text{für alle } x \in \mathbb{R}.$$

Dies ist die Verteilungsfunktion von $N(0, t)$. Definiere $W_t := N(0, t)$. Somit konvergiert für jedes feste t die Verteilung von $b^{(N)}(t)$ gegen die von W_t . \square

Satz 4.3.2 (Kovarianzstruktur der Brownschen Bewegung). Aus der Verteilungskonvergenz der einzelnen Zeitpunkte kann man noch keinen sinnvollen Grenzprozess folgern. Die Zeitpunkte W_t sind zwar normalverteilt, aber der Prozess könnte trotzdem sprunghaft sein. Nun wird die Verteilungskonvergenz der endlich-dimensionalen Vektoren gezeigt, beziehungsweise die Kovarianz der Realisierungen bei benachbarten Zeitpunkten untersucht. Sei $0 \leq t_1 \leq \dots \leq t_k$ eine Zerlegung. Dann gilt

$$(b^{(N)}(t_1), \dots, b^{(N)}(t_k)) \xrightarrow[N \rightarrow \infty]{d} (W_{t_1}, \dots, W_{t_k}).$$

*Beweis.*¹ Für jedes N ist der Vektor $(b^{(N)}(t_1), \dots, b^{(N)}(t_k))$ gemeinsam normalverteilt, denn $b^{(N)}(t)$ ist eine lineare Kombination der i.i.d. Standardnormalen (η_i) .

Es genügt, Mittelwerte und Kovarianzen zu kontrollieren: Die Mittelwerte sind Null. Für $s, t \geq 0$ mit $k = \lfloor Nt \rfloor$, $\alpha = Nt - k$, $\ell = \lfloor Ns \rfloor$, $\beta = Ns - \ell$ gilt mit $\xi_n = \sum_{i=1}^n \eta_i$:

$$b^{(N)}(t) = \frac{1}{\sqrt{N}}(\xi_k + \alpha\eta_{k+1}), \quad b^{(N)}(s) = \frac{1}{\sqrt{N}}(\xi_\ell + \beta\eta_{\ell+1}).$$

Unabhängigkeit der η_i ($E(\eta_i) = 0$, $V(\eta_i) = 1$) liefert

$$\begin{aligned} \text{Cov}(b^{(N)}(s), b^{(N)}(t)) &= E \left[b^{(N)}(s) b^{(N)}(t) \right] \cdot \underbrace{E(b^{(N)}(s)) \cdot b^{(N)}(t)}_{=0} \\ &= \frac{1}{N} E[(\xi_k + \alpha\eta_{k+1})(\xi_\ell + \alpha\eta_{\ell+1})] \\ &= \frac{1}{N} [E(\xi_\ell \xi_k) + \alpha E(\xi_\ell \eta_{k+1}) + \beta E(\eta_{\ell+1} \xi_k) + \alpha\beta E(\eta_{\ell+1} \eta_{k+1})] \end{aligned}$$

Da die Erwartungswerte der η s gleich Null sind und die η s unabhängig sind, können die Terme weiter reduziert werden.

$$\begin{aligned} E(\xi_\ell \xi_k) &= \sum_{i=1}^{\ell} \sum_{j=1}^k E(\eta_i \eta_j) \begin{cases} = E(\eta_i)(\eta_j) = 0, & i \neq j \\ = E(\eta_i^2) = 1, & i = j \end{cases} \\ &= \sum_{i=1}^{\ell} \sum_{j=1}^k \delta_{i,j} = \min(\ell, k) \end{aligned}$$

Analog folgt für die anderen Terme z. B.

$$\alpha E(\xi_\ell \eta_{k+1}) = \alpha \sum_{i=1}^{\ell} E(\eta_i \eta_{k+1}) = \alpha \mathbf{1}_{\{k+1 \leq \ell\}},$$

und so weiter. Insgesamt ergibt sich

$$\text{Cov}(b^{(N)}(s), b^{(N)}(t)) = \frac{1}{N} \left(\min(\ell, k) + \alpha \mathbf{1}_{\{k+1 \leq \ell\}} + \beta \mathbf{1}_{\{\ell+1 \leq k\}} + \alpha\beta \mathbf{1}_{\{\ell=k\}} \right)$$

¹angelehnt an Henze [9] S. 225f.

Im Grenzübergang gilt

$$\text{Cov}(b^{(N)}(s), b^{(N)}(t)) = \frac{\min(\ell, k)}{N} + O\left(\frac{1}{N}\right) \xrightarrow{N \rightarrow \infty} \min(s, t).$$

Folglich konvergiert die Kovarianzmatrix der Vektoren gegen $\Sigma = (\min(t_i, t_j))_{i,j}$. Mit der Cramér–Wold-Technik reicht es, lineare Formen zu betrachten. Sei also $a = (a_1, \dots, a_k)^\top \in \mathbb{R}^k$ und

$$Y_N := \sum_{i=1}^k a_i b^{(N)}(t_i), \quad Y := \sum_{i=1}^k a_i W_{t_i},$$

wobei

$$V(Y_N) = \sum_{i,j=1}^k a_i a_j \text{Cov}(b^{(N)}(t_i), b^{(N)}(t_j)),$$

und damit

$$V(Y_N) \xrightarrow{N \rightarrow \infty} \sum_{i,j=1}^k a_i a_j \min(t_i, t_j) =: \sigma^2(a).$$

Jedes Y_N ist eine lineare Kombination unabhängiger η_i . Mit dem Zentralen Grenzwertsatz (von Lindeberg–Feller) wird die Grenzverteilung von Y_N bestimmt. Da die Summanden zentriert und von endlicher Varianz sind, bleibt die Lindeberg-Bedingung zu zeigen. Zerlege

$$Y_N = \sum_{i=1}^k a_i b^{(N)}(t_i) = \sum_{j=1}^{M_N} c_{N,j} \eta_j,$$

wobei $M_N = \lfloor N t_k \rfloor + 1$ und die Koeffizienten $c_{N,j}$ aus der Definition von $b^{(N)}(t_i)$ stammen. Aus der Darstellung folgt, dass $|c_{N,j}| \leq C/\sqrt{N}$ für ein von N unabhängiges $C > 0$ gilt. Damit erhält man für jedes $\varepsilon > 0$

$$\mathbf{1}_{\{|c_{N,j} \eta_j| > \varepsilon\}} \leq \mathbf{1}_{\{|\eta_j| > \varepsilon \sqrt{N}/C\}}.$$

Also

$$L_N(\varepsilon) = \sum_{j=1}^{M_N} E(c_{N,j}^2 \eta_j^2 \mathbf{1}_{\{|c_{N,j} \eta_j| > \varepsilon\}}) \leq \left(\sum_{j=1}^{M_N} c_{N,j}^2 \right) E\left[\eta_1^2 \mathbf{1}_{\{|\eta_1| > \varepsilon \sqrt{N}/C\}} \right].$$

Die Summe der Quadrate $\sum_j c_{N,j}^2$ ist beschränkt, da sie die Varianz von Y_N ist und gegen $\sigma^2(a)$ konvergiert. Andererseits gilt wegen $\eta_1 \in L^2$ und $\varepsilon \sqrt{N}/C \rightarrow \infty$

$$E\left[\eta_1^2 \mathbf{1}_{\{|\eta_1| > \varepsilon \sqrt{N}/C\}} \right] \xrightarrow{N \rightarrow \infty} 0.$$

Damit ist die Lindeberg-Bedingung erfüllt und es folgt

$$Y_N \xrightarrow{d} \mathcal{N}(0, \sigma^2(a)).$$

Damit ist Y_N zentriert normalverteilt. Insgesamt konvergiert für jedes a die Verteilung von $a^\top (b^{(N)}(t_1), \dots, b^{(N)}(t_k))$ gegen die von $a^\top (W_{t_1}, \dots, W_{t_k})$. Nach Cramér–Wold folgt die behauptete Verteilungskonvergenz des Vektors. \square

Bemerkung 4.3.3 (Interpretation der Kovarianzen). Um die Wichtigkeit der Kovarianz-Struktur zu verdeutlichen, wird in der folgenden Visualisierung gezeigt, wie ein stochastischer Prozess ohne Kovarianz der einzelnen Zufallsvariablen aussehen würde. In Blau werden die Verteilungen der Zufallsvariablen, und in Rot jeweils Realisierungen des Prozesses dargestellt. In der zweiten Grafik (mit Kovarianz-Struktur) kann man beobachten, dass die Verteilungen eine geringere Varianz haben, und jeweils um den Vorherigen Wert zentriert sind. Dies wird im Folgenden noch präzisiert. Das kann man als ein Stetigkeits-Kriterium betrachten. (Obwohl man nicht direkt auf die Stetigkeit der Pfade folgern kann)

Ausgehend von der Kovarianz-Struktur kann man die Stetigkeit mit dem Stetigkeitssatz von Kolmogorov (Behrends [2], 2013, S. 78f.) folgern. In dieser Arbeit wird jedoch ein anderer Beweisansatz genutzt.

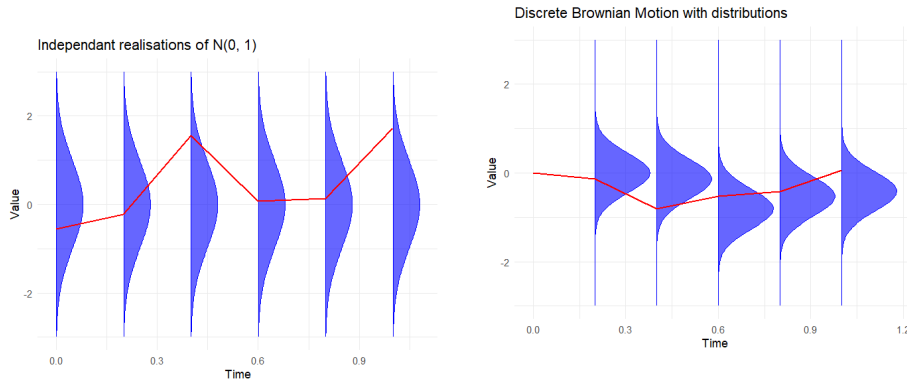


Abbildung 2: Folge von Normalverteilten Zufallsvariablen ohne Kovarianz

Abbildung 3: Visualisierung der Verteilung der diskreten Brownschen Bewegung unter Beachtung der Kovarianzstruktur

Satz 4.3.4 (Selbstähnlichkeit und bedingte Verteilung der Brownschen Bewegung). Für jedes $c > 0$ gilt

$$W_{ct}(\omega) \stackrel{d}{=} \sqrt{c} W_t(\omega) \quad \text{für alle } t \geq 0,$$

für alle $s < t$ gilt

$$W_t - W_s \sim N(0, t - s),$$

und die bedingte Verteilung von W_t gegeben W_s ist

$$W_t | W_s \sim N(W_s, t - s).$$

Beweis der ersten Behauptung (Selbstähnlichkeit).

Für $t \geq 0$ gilt $W_{ct} \sim N(0, ct)$ und $W_t \sim N(0, t)$. Somit gilt

$$\sqrt{c} W_t \sim N(0, ct).$$

Da beide Normalverteilungen denselben Mittelwert 0 und dieselbe Varianz ct haben, folgt

$$W_{ct} \stackrel{d}{=} \sqrt{c} W_t \quad \text{für jedes } t \geq 0.$$

□

Beweis der zweiten Behauptung (Inkremente).

Für $s < t$ gilt

$$W_t - W_s = \lim_{N \rightarrow \infty} (b^{(N)}(t) - b^{(N)}(s)),$$

wobei $b^{(N)}$ diskrete Approximierungen sind. Jedes Inkrement $b^{(N)}(t) - b^{(N)}(s)$ ist normalverteilt mit Erwartungswert 0 und Varianz $t - s$. Durch Grenzwertbildung folgt

$$W_t - W_s \sim N(0, t - s).$$

□

Beweis der dritten Behauptung (Bedingte Verteilung).

Betrachte $(W_s, W_t)^T$, das multivariat normal verteilt ist mit

$$E(W_s) = E(W_t) = 0, \quad \text{Cov}(W_s, W_t) = \min(s, t) = s.$$

Die Kovarianzmatrix lautet somit

$$\Sigma = \begin{pmatrix} s & s \\ s & t \end{pmatrix}.$$

Für die bedingte Verteilung ergibt die Standardformel der multivariaten Normalverteilung

$$\begin{aligned} E(W_t \mid W_s) &= E(W_t) + \text{Cov}(W_t, W_s) V(W_s)^{-1} (W_s - E(W_s)) \\ &= 0 + \frac{s}{s} W_s = W_s, \\ V(W_t \mid W_s) &= V(W_t) - \frac{\text{Cov}(W_t, W_s)^2}{V(W_s)} \\ &= t - \frac{s^2}{s} = t - s. \end{aligned}$$

Insgesamt folgt

$$W_t \mid W_s \sim N(W_s, t - s).$$

□

Korollar 4.3.5 (Martingal-Eigenschaft der Brownschen Bewegung).

Die Brownsche Bewegung W_t ist ein Martingal. *Beweis.* Aus dem vorigen Satz folgt

$$E(W_t \mid W_s = v) = v.$$

Damit ist W_t ein Martingal.

□

Satz 4.3.6 (Stetigkeit der Pfade). Der Pfad $t \mapsto W_t(\omega)$ ist fast sicher stetig. *Beweis.* Für den Beweis wird eine neue Funktionen-Folge $\hat{W}_t^{(n)}(\omega), n \in \mathbb{N}$ definiert, wobei $\hat{W}_t^{(n)}(\omega)$ die lineare Interpolation der Werte $W_{k/2^n}(\omega), k = 0, 1, 2, \dots$ ist. Ohne Beschränkung der Allgemeinheit wird das Intervall $[0, 1]$ betrachtet. Für $n \in \mathbb{N}$ und $k = 0, \dots, 2^n - 1$ setze

$$I_{n,k} := [k2^{-n}, (k+1)2^{-n}].$$

Da die $I_{n,k}$ eine Zerlegung von $[0, 1]$ bilden, ist

$$\begin{aligned} M_n &:= \sup_{t \in [0,1]} |\hat{W}_t^{(n+1)} - \hat{W}_t^{(n)}| \\ &= \max_{0 \leq k < 2^n} \sup_{t \in I_{n,k}} |\hat{W}_t^{(n+1)} - \hat{W}_t^{(n)}| = \max_{0 \leq k < 2^n} |Z_{n,k}| \end{aligned}$$

Für die unabhängigen Inkremente $Z_{n,k}$ mit

$$Z_{n,k} \sim N(0, 2^{-(n+2)}).$$

Für eine Normalverteilung $Z \sim N(0, \sigma^2)$ und jedes $\varepsilon > 0$ gilt (siehe z. B. Boucheron 2013 [4] S. 2)

$$P(|Z| > \varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right).$$

Mit der Schranke folgt

$$P(M_n > \varepsilon) \leq \sum_{k=0}^{2^n-1} P(|Z_{n,k}| > \varepsilon) \leq 2^n \cdot 2 \exp\left(-\frac{\varepsilon^2}{2\sigma_n^2}\right) = 2^{n+1} \exp\left(-\frac{\varepsilon^2}{2\sigma_n^2}\right).$$

Man wählt nun eine spezifische Folge ε_n so, dass sich aus der rechten Seite eine geometrische Reihe ergibt. Setze

$$\varepsilon_n := \sqrt{2(n+1)} \sigma_n.$$

Dann gilt

$$\frac{\varepsilon_n^2}{2\sigma_n^2} = n+1,$$

und damit

$$P(M_n > \varepsilon_n) \leq 2^{n+1} e^{-(n+1)} = (2e^{-1})^{n+1}.$$

Da $2e^{-1} < 1$ ist, ist die Folge auf der rechten Seite geometrisch und insbesondere summierbar. Folglich

$$\sum_{n=1}^{\infty} P(M_n > \varepsilon_n) < \infty.$$

Mit dem Reihenkriterium für fast sichere Konvergenz folgt, dass $M_n \rightarrow 0$ fast sicher. Für $m > n \geq N$ beliebig gilt

$$\sup_t |\hat{W}_t^{(m)} - \hat{W}_t^{(n)}| \leq \sum_{k=n}^{m-1} M_k \xrightarrow[N \rightarrow \infty]{} 0$$

Weil jedes $(\hat{W}^{(n)})_{n \in \mathbb{N}}$ eine stückweise lineare Funktion ist, folgt dass $(\hat{W}^{(n)})_{n \in \mathbb{N}}$ fast sicher eine Cauchy-Folge in $\|\cdot\|_\infty$ ist. $(\hat{W}^{(n)})_{n \in \mathbb{N}}$ konvergiert also gleichmäßig gegen den Grenzpfad \widetilde{W} , der stetig ist. \square

Achtung: Es wurde nicht gezeigt, der Prozess $W_t^{(n)} \rightarrow W_t$ gleichmäßig konvergiert. Für eine beliebige aber feste Realisierung $t \mapsto W_t(\omega)$ wurde eine (reellwertige) Folge von stetigen Funktionen konstruiert, die gegen $t \mapsto W_t$ gleichmäßig konvergiert. Da es sich um eine Folge stetiger Funktionen handelt, gilt die Stetigkeit unter gleichmäßiger Konvergenz auch gegen die Grenzfunktion $t \mapsto W_t$.

5 Die geometrische Brownsche Bewegung

Die geometrische Brownsche Bewegung ist eine Erweiterung der Brownschen Bewegung. Sie eignet zur Modellierung von Aktienkursen, da sie im Gegensatz zur klassischen Brownschen Bewegung stets positive Werte annimmt.

5.1 Die geometrische Brownschen Bewegung

Das *Binomialmodell* beschreibt den Aktienkurs S_t in diskreter Zeit: In jedem Zeitschritt ändert sich der Kurs multiplikativ um einen Zufallsfaktor. Es gilt

$$S_{k+1} = S_k (1 + X_{k+1}), \quad k = 0, 1, \dots, n-1,$$

für eine Zufallsvariable X_{k+1} die die relative Kursänderung im Schritt $k+1$ repräsentiert. Um eine kontinuierliche Zeitentwicklung zu modellieren, setzt man²

$$X_{k+1} = \mu \Delta t + \sigma \sqrt{\Delta t} \varepsilon_{k+1},$$

mit μ als Erwartungswert der Rendite (oder auch Drift), σ als Volatilität, und ε_{k+1} unabhängig, identisch verteilten Zufallsvariablen mit Erwartungswert 0 und Varianz 1. (An dieser Stelle ist unwichtig, wie die ε_{k+1} verteilt sind, es reicht, dass sie diese Momente besitzen. Die Information über die Verteilung wird in einem Grenzübergang verloren gehen.) Das bildet die Modellannahme für den Rest dieser Arbeit.

Im Folgenden wird eine explizite Formel für S_t bewiesen, nämlich

$$S_T = S_0 \exp\left((\mu - \frac{1}{2}\sigma^2)T + \sigma W_T\right),$$

wobei W_T eine Brownsche Bewegung ist und $T = n \cdot \Delta t$. *Beweis.* Man betrachtet den Logarithmus der S_k : Nach n Schritten ist der Aktienkurs

$$S_n = S_0 \prod_{j=1}^n (1 + X_j).$$

²Dies ist ein Schritt des Euler-Maruyama-Verfahren zur numerischen Lösung der stochastischen Differentialgleichung der geometrischen Brownschen Bewegung. Vgl. Kapitel 8 oder Bärwolff und Tischendorf [1].

Durch den Logarithmus erhält man

$$\log S_n = \log S_0 + \sum_{j=1}^n \log(1 + X_j). \quad (5.1.1)$$

Als nächstes wird die Taylor-Entwicklung der Terme $\log(1 + X_j)$ betrachtet. Die k -te Ableitung lautet

$$\log(1 + x)^{(k)} = (-1)^{k+1} \frac{(k-1)!}{(1+x)^k}.$$

Setzt man diese in die Taylor-Formel ein ergibt sich

$$\log(1 + x) = \sum_{k=0}^{\infty} \frac{(\log(1 + \cdot)^{(k)})(0)}{k!} (x - 0)^k = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$$

Da $X_j \in O(\sqrt{\Delta t})$ ist, reicht die Taylor-Entwicklung bis zum quadratischen Term: Es gilt

$$X_j^k = (\mu \Delta t + \sigma \sqrt{\Delta t} \varepsilon_j)^k.$$

Durch den Binomischen Lehrsatz ergibt sich:

$$X_j^k = \sum_{m=0}^k \binom{k}{m} (\mu \Delta t)^{k-m} (\sigma \sqrt{\Delta t} \varepsilon_j)^m.$$

Die Terme enthalten Potenzen von Δt in der Form $(\Delta t)^{(k-m)+m/2} = (\Delta t)^{k-m/2}$. Für $k \geq 3$ (und $m = 0, \dots, k$) ist $(\Delta t)^{k-m/2}$ von höherer Ordnung als Δt oder $\sqrt{\Delta t}$ und verschwindet daher im Grenzübergang $\Delta t \rightarrow 0$ bzw. $n \rightarrow \infty$. Der Grenzübergang von diskreter zu kontinuierlicher Zeit führt eben dazu, dass n und Δt gleichzeitig gegen ∞ bzw. 0 gehen, wobei $n\Delta t = T$ konstant bleibt. Daher verschwinden in der Summe in Formel 5.1.1 genau die Terme, in denen Δt einen Exponenten größer 1 hat.

$$\log(1 + X_j) \approx X_j - \frac{1}{2} X_j^2.$$

Der quadratische Term wird nun ausmultipliziert und in der selben Weise abgeschätzt:

$$\frac{1}{2} X_j^2 \approx \frac{1}{2} \sigma^2 \Delta t \varepsilon_j^2.$$

Hier verschwinden die Terme $\mu(\Delta t)^2 \in O((\Delta t)^2)$ und $2\mu(\Delta t)\sigma\sqrt{\Delta t}\varepsilon_{k+1} \in O((\Delta t)^{3/2})$ wieder im Limes. Zwischenfazit:

$$\begin{aligned} \log S_n &\approx \log S_0 + \sum_{j=1}^n \left(\underbrace{\mu \Delta t + \sigma \sqrt{\Delta t} \varepsilon_j}_{X_j} - \underbrace{\frac{1}{2} \sigma^2 \Delta t \varepsilon_j^2}_{-\frac{1}{2} X_j^2} \right) \\ &= \log S_0 + \mu T + \sigma \sqrt{\Delta t} \sum_{j=1}^n \varepsilon_j - \frac{1}{2} \sigma^2 \sum_{j=1}^n \Delta t \varepsilon_j^2 \end{aligned}$$

Im Folgenden wird der Grenzübergang $n \rightarrow \infty$ bzw. $\Delta t \rightarrow 0$ durchgeführt. Die erste Summe konvergiert nach dem Zentralen Grenzwertsatz (Verteilungskonvergenz):

$$\sigma\sqrt{\Delta t} \sum_{j=1}^n \varepsilon_j = \sigma \frac{\sqrt{T}}{\sqrt{N}} \sum_{j=1}^n \varepsilon_j \xrightarrow[n \rightarrow \infty]{d} \xi \sim N(0, \sigma^2 T),$$

also gegen eine Zufallsvariable ξ die normalverteilt ist, mit Varianz $\sigma^2 T$. $\xi = \sigma W_T$ ist eine Lösung, weil $W_T \sim N(0, T)$ eine Brownsche Bewegung zur Zeit T ist. Da $E(\varepsilon_j^2) = 1$ gilt, folgt im Grenzübergang für die zweite Summe nach dem Gesetz der großen Zahlen

$$\frac{1}{2}\sigma^2 \sum_{j=1}^n \Delta t \varepsilon_j^2 \xrightarrow[n \rightarrow \infty]{f.s.} \frac{1}{2}\sigma^2 T.$$

Damit ergibt sich im Grenzübergang $n \rightarrow \infty$:

$$\log S_T \stackrel{d}{=} \log S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W_T.$$

Exponentiell geschrieben erhält man die *Geometrische Brownsche Bewegung*:

$$S_T \stackrel{d}{=} S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W_T\right).$$

□

5.2 Die logarithmische Normalverteilung

Eine Zufallsvariable X heißt log-Normalverteilt mit Varianz σ^2 und Erwartungswert μ , falls die Zufallsvariable $Y := \log(X)$ normalverteilt ist mit $Y \sim N(\mu, \sigma^2)$. S_T ist somit log-normalverteilt:

$$\log S_T \stackrel{d}{=} \log S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W_T.$$

ergibt

$$\log S_T \sim N\left(\log S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)T, \sigma^2 T\right)$$

Erwartungswert und Varianz eines Kurses zum Zeitpunkt T sind damit

$$E(S_T) = \log S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)T, \quad V(S_T) = \sigma^2 T.$$

6 Anwendungen auf Zeitreihen

6.1 Kalibrierung

Aus einem Datensatz lassen sich die Parameter μ und σ der geometrischen Brownschen Bewegung schätzen. Für reale Werte ist $\Delta t > 0$ und n ist die

(endliche) Anzahl von Datenpunkten. Zur Schätzung von μ und σ werden die log-Rendite

$$r_j := \log S_j - \log S_{j-1} = \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma(W_j - W_{j-1})$$

genutzt. Da $W_j - W_{j-1} \sim N(0, \Delta t)$ folgt

$$r_j \sim N\left(\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t, \sqrt{\sigma}\Delta t\right).$$

Man berechnet also die log-Rendite \hat{r}_j des Datensatzes, und davon den empirischen Erwartungswert m (den Durchschnitt) und die empirische Varianz s^2 . Dann folgt

$$\sigma \approx s, \quad \mu \approx m + \frac{1}{2}s^2.$$

Die Schätzung der Parameter kann in R wie folgt durchgeführt werden:

```
1 log_returns <- diff(log(dax$Price)) # tägliche Werte
2 sigma <- sd(log_returns)
3 mu <- mean(log_returns) + 0.5 * sigma^2
```

6.2 Bootstrap-Verfahren zur Kalibrierung

Das Bootstrap-Verfahren ermöglicht eine probabilistische Schätzung der Parameter μ und σ . Falls die vorliegenden Daten z. B. Lückenhaft sind, kann dies detektiert werden. Durch wiederholte Auswahl (Sampling) von Teildaten und Schätzung der Parameter, wobei die verschiedenen Schätzungen anschließend verglichen werden kann man Konfidenzintervalle für μ und σ erzeugen. Wenn sich die Parameterschätzungen stark unterscheiden kann man z. B. schließen, dass die Daten unsauber sind, oder nicht der GBM entsprechen. Es folgt eine Implementierung in R:

```
1 log_returns <- diff(log(dax$Price))
2 boot_res <- replicate(1000, {
3   sample_ret <- sample(log_returns, length(log_returns) / 2, replace
4     = TRUE)
5   c(mu = mean(sample_ret) * 252 - 0.5 * (sd(sample_ret) * sqrt(252))
6     ^ 2,
7     sigma = sd(sample_ret) * sqrt(252))
8 })
9 boot_df <- as.data.frame(t(boot_res))
10 mu_ci <- quantile(boot_df$mu, c(0.025, 0.975))
11 sigma_ci <- quantile(boot_df$sigma, c(0.025, 0.975))
```

Die Bootstrap-Methode liefert für die DAX-Daten zum 5%-Niveau:

$$\mu = 0.051 \in [-0.055, 0.148], \quad \sigma = 0.218 \in [0.212, 0.225].$$

6.3 Berechnung von Konfidenzintervallen

Da S_T log-normalverteilt ist, reicht es ein Konfidenzintervall für die log-Normalverteilung zu berechnen. Sei $X \sim N(\mu, \sigma^2)$ eine normalverteilte Zufallsvariable. Dann ist $Y := e^X$ log-normalverteilt mit Parametern μ und σ^2 . Ein zweiseitiges Konfidenzintervall für X mit Konfidenzniveau $1 - \alpha$ ist

$$[\mu - z_{\alpha/2}\sigma, \mu + z_{\alpha/2}\sigma],$$

wobei $z_{\alpha/2}$ das $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung ist. Exponentiell transformiert ergibt sich das Konfidenzintervall für Y :

$$[e^{\mu - z_{\alpha/2}\sigma}, e^{\mu + z_{\alpha/2}\sigma}].$$

Beispiel 6.3.1 (Konfidenzintervall für den DAX). Im folgenden R-Programm (Ausschnitt) wird ein 95%-Konfidenzintervall für den DAX in einem Jahr von heute (252 Handelstage) berechnet. Dazu werden die Parameter μ und σ wie oben aus den täglichen log-Renditen geschätzt.

```
1 alpha = 0.05
2 T <- 252
3 z <- qnorm(c(1 - alpha/2, alpha/2))
4 ci <- S0 * exp((mu - 0.5 * sigma^2) * T + z * sigma * sqrt(T))
```

Hier ist S_0 der heutige Kurswert des DAX. Das Konfidenzintervall lautet in diesem Fall:

$$[17217, 40097].$$

Beispiel 6.3.2 (Konfidenzband für den DAX). Im folgenden R-Programm (Ausschnitt) wird ein 95%-Konfidenzband für den DAX im nächsten Jahr (252 Handelstage) berechnet.

```
1 alpha = 0.05
2 n <- 252
3 last_date <- max(dax$Date)
4 future_dates <- last_date + 1:n
5
6 q_low <- S0 * exp((mu - 0.5 * sigma^2) * (1:n) + qnorm(1 - alpha/2) *
   sigma * sqrt((1:n)))
7 q_hi  <- S0 * exp((mu - 0.5 * sigma^2) * (1:n) + qnorm(alpha/2) *
   sigma * sqrt((1:n)))
8 q_med <- S0 * exp((mu - 0.5 * sigma^2) * (1:n))
9
10 band <- data.frame(Date = future_dates, low = q_low, mid = q_med, hi
   = q_hi)
```

Es folgt eine Visualisierung des Konfidenzbandes im Anschluss an die historischen Daten.



Abbildung 4: DAX mit 95%-Konfidenzband für das nächste Jahr

6.4 Back-Tests

Ein Backtest überprüft im Nachhinein, wie gut ein Modell mit historisch tatsächlich verfügbaren Daten funktioniert hätte.

Beispiel 6.4.1 (Back-Test der Kalibrierung auf den DAX). Im Fall der DAX-Zeitreihe werden die Parameter μ und σ auf die Daten von 1990 bis 2008 geschätzt. Nun wird ein Backtest durchgeführt, indem die Parameter genutzt werden, um die DAX-Entwicklung von 2008 bis 2025 zu simulieren. Insgesamt wird also ein 50/50-Mix aus historischen Daten und simulierten Daten betrachtet. Außerdem werden Konfidenzintervalle zum Niveau 50% berechnet. Die folgende Abbildung 5 zeigt die Ergebnisse.

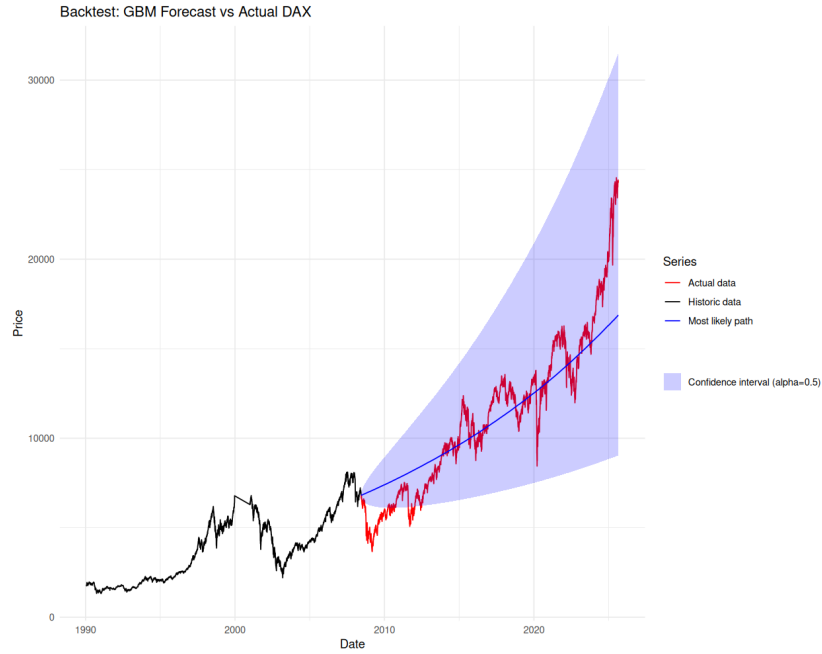


Abbildung 5: DAX Backtest mit 50% Konfidenzintervall und erwarteter Entwicklung

In diesem Fall fallen 86% der historischen Daten in das 50%-Konfidenzintervall, was für die Einfachheit des Modells ein gutes Ergebnis ist. Diese Metrik nennt man *Überdeckungswahrscheinlichkeit*.

Lemma 6.4.2 (Weitere Metriken für Zeitreihenschätzungen). Botchkarev [3] nennt u. a. folgende Metriken, um die Richtigkeit von Vorhersagen zu bewerten:

- Mean Squared Error (MSE):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- Normalized Root Mean Squared Error (NRMSE) :

$$\text{NRMSE} = \frac{\sqrt{\text{MSE}}}{y_{\max} - y_{\min}}$$

- Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Hierbei sind y_i die tatsächlichen Werte und \hat{y}_i die vorhergesagten Werte. Der MSE misst den durchschnittlichen quadratischen Fehler, der bei großen Werten der Zeitreihe schwer zu interpretieren ist. Der NRMSE normiert den MSE auf den Bereich der Zeitreihe, wodurch er vergleichbarer wird. Der MAPE gibt den durchschnittlichen prozentualen Fehler an, und ist somit ebenfalls anschaulich interpretierbar.

Beispiel 6.4.3 (Metriken des DAX-Backtests). Im obigen Backtest des DAX ergeben sich folgende Metriken:

Metrik	Wert	Einheit
MSE	3489387	Punkte ²
MAPE	13.91685	%
NRMSE	0.1659164	-

Tabelle 1: Fehlermaße der Prognose

Beispiel 6.4.4 (Sequenzieller Backtest). Man kann Backtests auch sequenziell durchführen: Für jedes Jahr wird aus den vorherigen Jahren eine GBM kalibriert, die dann mit dem tatsächlichen Wert verglichen wird. Dies wird in Abbildung 6 gezeigt.

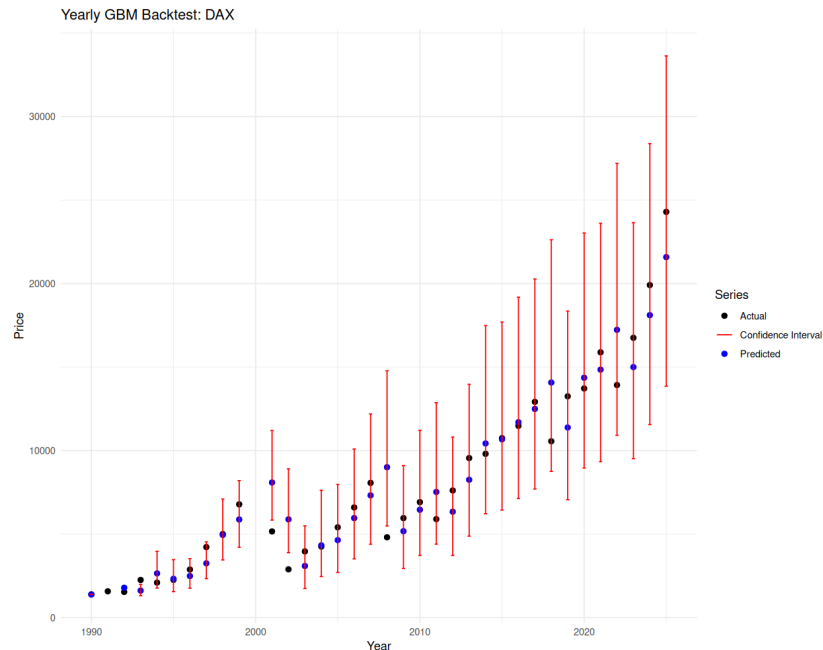


Abbildung 6: Sequenzieller Backtest für den DAX

6.5 Monte-Carlo-Simulation

Genauso wie man eine Brownsche Bewegung mit summierten Normalverteilungen simuliert, wird eine geometrische Brownsche Bewegung durch exponentiell transformierte summierte Normalverteilungen simuliert.

Beispiel 6.5.1 (Monte-Carlo-Simulation des DAX). Theoretisch liegt das stetige Modell zugrunde, aber in der Praxis wird eine diskrete Approximation verwendet, hier mit täglichen Schritten. Der folgende R-Code (Ausschnitt) simuliert 1000 Pfade der geometrischen Brownschen Bewegung mit den oben geschätzten Parametern μ und σ für den DAX, wieder für das nächste Jahr (252 Handelstage).

```
1 n <- 252
2 paths <- 1000
3 S0 <- tail(dax$Price, 1)
4
5 simulations <- replicate(paths, {
6   W <- c(0, cumsum(rnorm(n, 0, 1)))
7   S0 * exp((mu - 0.5 * sigma^2) * c(0, (1:n)) + sigma * W)
8 })
```

Es folgt eine Visualisierung der Simulation im Anschluss an die historischen Daten.

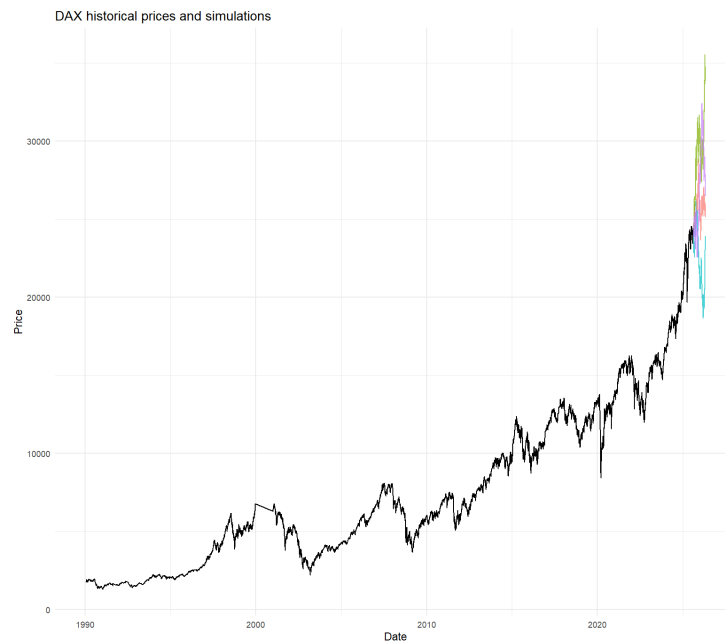


Abbildung 7: DAX mit 4 simulierten Pfaden für das nächste Jahr

Beispiel 6.5.2 (Vergleich von Konfidenzintervall und Monte-Carlo-Simulation). Man kann das Konfidenzintervall aus dem vorherigen Beispiel mit den quantilen der Monte-Carlo-Simulation vergleichen. Man erkennt, dass das Konfidenzintervall aus der Formel mit steigender Simulationsanzahl immer besser durch die Quantile der Simulation approximiert wird.

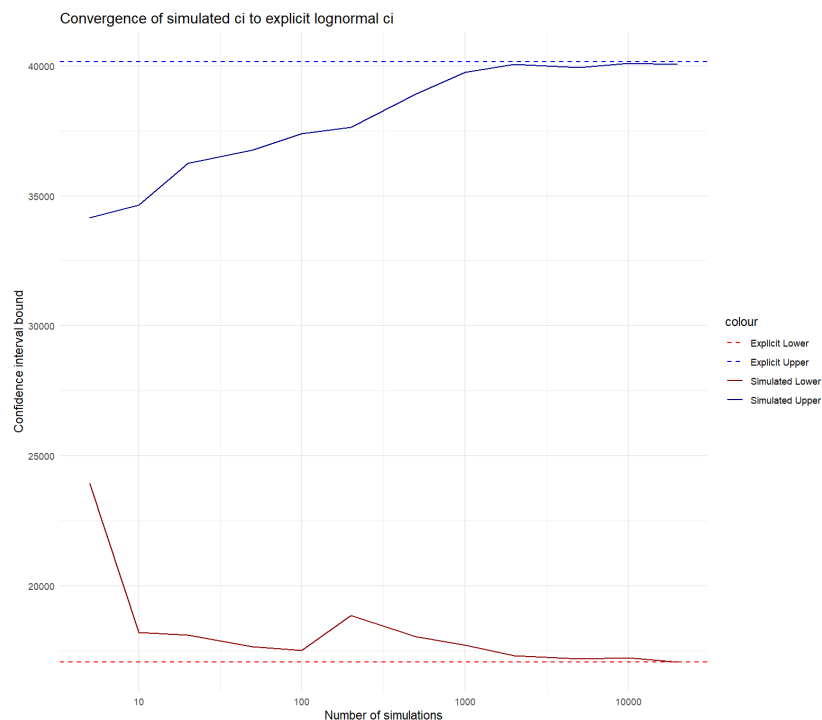


Abbildung 8: Vergleich von Konfidenzintervall und Monte-Carlo-Simulation für verschiedene Simulationsanzahlen

7 Aktienoptionen

7.1 Finanzmathematische Grundlagen

Aktienoptionen sind Finanzderivate, die dem Inhaber das Recht, aber nicht die Pflicht geben, eine Aktie zu einem vorher festgelegten Preis (dem Ausübungspreis) zu kaufen (Call-Option) oder zu verkaufen (Put-Option). Europäische Optionen können nur am Fälligkeitstag ausgeübt werden, während amerikanische Optionen während der gesamten Laufzeit bis zum Verfallsdatum ausgeübt werden können. Den fairen Preis einer Option zu bestimmen ist eine Herausforderung, die man mit Hilfe der geometrischen Brownschen Bewegung lösen kann. Im Folgenden wird dazu das Black-Scholes-Modell vorgestellt.

Beispiel 7.1.1. *Beispiel Call-Option.* Ein Investor erwirbt eine europäische Call-Option auf die Aktie der Firma X mit einem Ausübungspreis von 50 €. Am Fälligkeitstag steht der Aktienkurs bei 60 €. Der Investor übt die Option aus, kauft die Aktie für 50 € und kann sie sofort für 60 € verkaufen. Sein Gewinn (ohne Berücksichtigung der Optionsprämie) beträgt 10 € pro Aktie.

Motivation. Der Investor spekuliert darauf, dass der Kurs der Aktie steigt und über den Ausübungspreis hinausgeht.

Beispiel Put-Option. Ein Landwirt sichert sich gegen fallende Weizenpreise ab und kauft eine europäische Put-Option mit einem Ausübungspreis von 200 €/Tonne. Am Fälligkeitstag liegt der Marktpreis bei 170 €/Tonne. Der Landwirt übt die Option aus und verkauft seinen Weizen zum höheren Preis von 200 €/Tonne, obwohl der Marktpreis niedriger ist. Sein Vorteil beträgt 30 € pro Tonne (abzüglich der Optionsprämie).

7.2 Bewertung von Aktienoptionen im Binomialmodell mit dem Risikoneutralen Maß

Das Binomialmodell wird erweitert. Zusätzlich zur Aktie und deren Kurs gibt es nun die Bank, die einen risikofreien Zinssatz r anbietet. Zusätzlich wird ein neues Wahrscheinlichkeitsmaß Q eingeführt, das sogenannte risikoneutrale Maß.

Definition 7.2.1 (Arbitrage-Prinzip). Durch Arbitrage erzielt man Gewinne, in dem man Güter an einem Markt kauft, und dann fast gleichzeitig (teurer) an einem anderen verkauft. Die Möglichkeit für Arbitrage existiert nur kurzfristig, da die Preise sich im Weltmarkt ausgleichen. Im vorliegenden Modell wird Arbitrage daher ausgeschlossen. Es wird angenommen, dass man keine risikolosen Gewinne erzielen kann, ohne Kapital zu investieren.

Definition 7.2.2 (Europäische Optionen). Europäische Optionen geben dem Inhaber das Recht, einen Basiswert zu einem festgelegten Preis (dem Ausübungspreis) nur am Fälligkeitstag zu kaufen (Call-Option) oder zu verkaufen (Put-Option). Der Wert einer europäischen Call-Option am Endzeitpunkt T ist durch die Auszahlungsfunktion

$$f(S_T) = \max(S_T - K, 0) = (S_T - K)^+$$

gegeben, wobei K der Ausübungspreis ist. $S_T - K$ ist der Gewinn der erzielt wird, wenn der Inhaber Aktien zu einem niedrigeren Kurs (K) erwirbt, und dann zum echten Kurs (S_T) verkauft. Die Auszahlungsfunktion kann keine negativen Werte annehmen, da der Inhaber wohl nicht auf sein Recht bestehen würde, Aktien zu einem erhöhten Kurs zu kaufen. Für eine Put-Option ist die Auszahlungsfunktion

$$\tilde{f}(S_T) = \max(K - S_T, 0) = (K - S_T)^+.$$

Das entspricht der *Call-Put-Parität* $f(S_T) - \tilde{f}(S_T) = S_T - K$.

Definition 7.2.3 (Hedging). Ein Anbieter von Optionen kann sich absichern, indem er ein sogenanntes Hedge-Portfolio bildet. Dieses besteht aus einer dynamischen Position im Basiswert sowie einer Position im risikofreien Konto. Ziel ist es, die Auszahlungsstruktur der Option durch das Portfolio exakt zu replizieren.

Formal bedeutet Hedging: Es existieren Prozesse Δ_t (Anzahl der gehaltenen Aktien) und B_t (Bestand im Bankkonto), so dass für alle Zeiten t gilt

$$V_t = \Delta_t S_t + B_t,$$

wobei V_t der Wert des Hedge-Portfolios ist. Ist das Portfolio selbstfinanzierend, d. h. Änderungen in V_t entstehen ausschließlich durch Änderungen in S_t und nicht durch externe Ein- oder Auszahlungen, und gilt zudem $V_T = H(S_T)$ mit der Optionsauszahlung $H(S_T)$, so spricht man von einer perfekten Replikation. Die Grundidee wird am Beispiel einer europäischen Call-Option verdeutlicht: Steigt der Wert der Aktie über den Ausübungspreis K und der Optionsanbieter ist verpflichtet, die Aktie zum Preis K zu liefern, so können dafür die im Hedge-Portfolio gehaltenen Aktien verwendet werden. Auf diese Weise wird der potenzielle Verlust ausgeglichen.

Lemma 7.2.4 (Risikoneutrale Wahrscheinlichkeit für einen Schritt). Ein Betrag B_0 der aus dem Bankkonto angelegt wird steigt durch den risikofreien Zinssatz r kontinuierlich durch

$$B(t) = e^{rt} B_0.$$

Daher ist ein Gewinn in der Zukunft (Δt) weniger wert als ein Gewinn jetzt, da man den jetzigen Gewinn G anlegen könnte, und dann $Ge^{r\Delta t}$ hätte.

Da dies berücksichtigt werden muss, werden mögliche Aktiengewinne mit dem Faktor $e^{-r\Delta t}$ verkleinert (diskontiert).

Der diskontierte Aktienkurs zum Zeitpunkt n ist $e^{-rn\Delta t} S_n$. Im Folgenden wird berechnet, für welche Wahrscheinlichkeit der neue diskontierte Prozess ein Martingal ist. Das Martingal-Kriterium lautet: $E(X_{n+1} \mid X_n = l) = l$, wobei $X_n = e^{-rn\Delta t} S_n$. Einsetzen von $X_{n+1} = e^{-r(n+1)\Delta t} S_{n+1}$ und $X_n = e^{-rn\Delta t} S_n$ ergibt

$$E(e^{-r(n+1)\Delta t} S_{n+1} \mid S_n = v) = e^{-rn\Delta t} v$$

Aus der Linearität des Erwartungswerts folgt

$$E(e^{-r\Delta t} S_{n+1} \mid S_n = v) = v.$$

Setzt man die möglichen Werte für S_{n+1} ein ($S_{n+1} = uS_n$ mit Wahrscheinlichkeit q , $S_{n+1} = dS_n$ mit Wahrscheinlichkeit $1 - q$) folgt

$$e^{-r\Delta t} (quS_n + (1 - q)dS_n) = v.$$

Teilt man durch $S_n = v$ ($v > 0$) ergibt sich

$$e^{-r\Delta t} (qu + (1 - q)d) = 1.$$

Das ist äquivalent zu

$$q = \frac{e^{r\Delta t} - d}{u - d}.$$

Zum Vergleich: im klassischen Binomialmodell ist

$$p = \frac{1 - d}{u - d}$$

die "risikoneutrale Wahrscheinlichkeit". Anders als zuvor wurde aber lediglich eine schwächere Aussage gezeigt (nur ein Schritt). Das wird am Folgenden jedoch nachgezogen.

Definition 7.2.5 (Risikoneutrales Maß). Die soeben berechnete Wahrscheinlichkeit q gehört zum risikoneutralen Wahrscheinlichkeitsmaß Q . Bisher wurde mit dem ursprünglichen Maß P gerechnet. Ab jetzt wird mit $E(X)$, $V(X)$ etc der Erwartungswert und die Varianz unter P bezeichnet, während $E_Q(X)$, $V_Q(X)$ etc den Erwartungswert und die Varianz unter Q bezeichnen. An Q werden folgende Anforderungen gestellt (vgl. Hull, 2022 [10], S. 270):

1. Q ist ein Wahrscheinlichkeitsmaß
2. Q ist äquivalent zu P , das heißt $P(A) = 0 \iff Q(A) = 0$ für alle Ereignisse A (lediglich die Wahrscheinlichkeiten ändern sich, nicht die möglichen Ereignisse)
3. Unter Q ist der diskontierte Aktienkurs ein Martingal, also $E_Q(e^{-r\Delta t} S_{n+1} | S_n = v) = v$ für alle n und $v > 0$.

Im Fall des Binomialmodells handelt es sich um ein diskretes Wahrscheinlichkeitsmaß, wobei die Werte $Q(S_{t+\Delta t} = u \cdot S_t) = q$ und $Q(S_{t+\Delta t} = d \cdot S_t) = 1 - q$ bekannt sind. Oft können Maße in der Stochastik nur indirekt über ihre Eigenschaften definiert werden, das gilt auch hier. Die Untersuchung der Existenz und Eindeutigkeit eines solchen Maßes überschreitet den Rahmen dieser Arbeit; Es sei auf Shreve [14], (2004, Kap. 5) verwiesen.

Lemma 7.2.6 (Diskontierter Aktienkurs im Binomialmodell ist unter Q ein Martingal). Mit der Schritt Wahrscheinlichkeit aus dem vorherigen Lemma wird nachgewiesen, dass der diskontierte Aktienkurs unter dem risikoneutralen Maß Q ein Martingal ist. Dies ergibt sich zwar aus den Anforderungen an Q , kann aber mit der Ein-Schritt-Übergangswahrscheinlichkeit verifiziert werden. Setze $X_n := e^{-rn\Delta t} S_n$. Unter dem risikoneutralen Maß Q , das heißt für $q = \frac{e^{r\Delta t} - d}{u - d}$ ist $(X_n)_{n \in \mathbb{N}_0}$ ein Martingal.

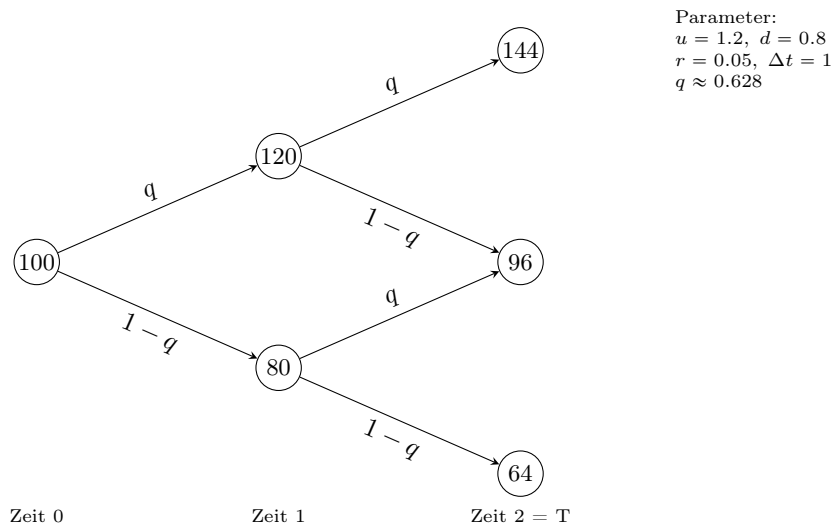
Beweis. Schreibe $S_{n+1} = S_n \xi_{n+1}$ mit $Q(\xi_{n+1} = u) = q$ und $Q(\xi_{n+1} = d) = 1 - q$. Für $s < t$ beliebig und wegen Unabhängigkeit der Schritte:

$$X_t = e^{-rt\Delta t} S_t = e^{-rt\Delta t} S_s \prod_{i=s+1}^t \xi_i,$$

$$\begin{aligned} E_Q(X_t \mid S_s = v) &= e^{-rt\Delta t} v \prod_{i=s+1}^t E_Q(\xi_i) = e^{-rt\Delta t} v (qu + (1-q)d)^{t-s} \\ &= e^{-rt\Delta t} v (e^{r\Delta t})^{t-s} = e^{-rs\Delta t} v = X_s. \end{aligned}$$

$$\begin{aligned} C_n &= e^{-r\Delta t} (qC_{n+1}^{\text{up}} + (1-q)C_{n+1}^{\text{down}}) \\ &= E_Q(e^{-r\Delta t}C_{n+1} \mid S_n) \end{aligned}$$

Beispiel 7.2.8. Seien $S_0 = 100$, $u = 1.2$, $d = 0.8$, $r = 0.05$, $\Delta t = 1$ Jahr, $K = 100$ und $T = 2$ Jahre. Die risikoneutrale Wahrscheinlichkeit ist $q = \frac{e^{0.05} - 0.8}{1.2 - 0.8} \approx 0.628$. Die möglichen Aktienkurse zum Zeitpunkt T sind:



Die Auszahlungen der Call-Option zum Zeitpunkt T sind:

$$\begin{aligned} C_{T,\text{up}} &= \max(144 - 100, 0) = 44, \\ C_{T,\text{mid}} &= \max(96 - 100, 0) = 0, \\ C_{T,\text{down}} &= \max(64 - 100, 0) = 0. \end{aligned}$$

Rückwärtsinduktion:

$$\begin{aligned} C_{1,\text{up}} &= e^{-0.05} (q \cdot 44 + (1 - q) \cdot 0) \approx 26.28, \\ C_{1,\text{down}} &= e^{-0.05} (q \cdot 0 + (1 - q) \cdot 0) = 0, \\ C_0 &= e^{-0.05} (q \cdot 23.99 + (1 - q) \cdot 0) \approx 15.7. \end{aligned}$$

Der faire Preis der Call-Option ist also etwa 15.7 €.

7.3 Das Black-Scholes-Modell als Grenzfall des Binomialmodells

Ziel ist es zu zeigen, dass die fairen Optionspreise beim Verfeinern der Zeitdiskretisierung gegen die geometrische Brownsche Bewegung konvergieren. Dieses Modell wird auch als Black-Scholes-Modell bezeichnet. Hier wird der Grenzübergang im risikoneutralen Maß durchgeführt, die Modellannahme aus Kapitel 5 wird übernommen:

$$S_t = S_0 \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

und durch Zeitdiskretisierung:

$$S_{t+\Delta t} = S_t \cdot e^{(\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma W_{t+\Delta t}}$$

Die risikoneutrale Wahrscheinlichkeit im Grenzübergang Zur Vereinfachung der zeitdiskreten Sprunggrößen u und d wird der Drift μ vernachlässigt. Das wird durch ein Arbitrage-Argument gerechtfertigt (vgl. Behrends[2], S. 133): Der deterministische Teil der Preisbewegung muss unter dem risikoneutralen Maß verschwinden, sonst gäbe es eine risikolose Gewinnmöglichkeit (Arbitrage). Für einen Zeitschritt Δt folgen damit die Sprunggrößen

$$u = e^{\sigma\sqrt{\Delta t}}, \quad d = e^{-\sigma\sqrt{\Delta t}}$$

und den risikoneutralen Schritt $q(\Delta t)$ aus dem Martingal-Kriterium

$$q(\Delta t) = \frac{e^{r\Delta t} - d}{u - d} = \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

Durch eine Taylor-Entwicklung um 0 werden die Funktionen wieder vereinfacht. Analog zum Beweis aus Kapitel 5 sollen höhere Potenzen von Δt im Grenzwert verschwinden. Da es sich um einen Quotienten handelt, kann man die Abschätzung aber nicht direkt anwenden. Stattdessen werden Zähler und Nenner separat entwickelt:

$$e^{r\Delta t} = 1 + r\Delta t + o(\Delta t), \quad e^{\pm\sigma\sqrt{\Delta t}} = 1 \pm \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t).$$

Damit

$$\begin{aligned}\text{Zähler} &= e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}} = (1 + r\Delta t + o(\Delta t)) - (1 - \sigma\sqrt{\Delta t} + \tfrac{1}{2}\sigma^2\Delta t + o(\Delta t)) \\ &= \sigma\sqrt{\Delta t} + (r - \tfrac{1}{2}\sigma^2)\Delta t + o(\Delta t), \\ \text{Nenner} &= e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}} = (1 + \sigma\sqrt{\Delta t} + \tfrac{1}{2}\sigma^2\Delta t + o(\Delta t)) \\ &\quad - (1 - \sigma\sqrt{\Delta t} + \tfrac{1}{2}\sigma^2\Delta t + o(\Delta t)) = 2\sigma\sqrt{\Delta t} + o(\sqrt{\Delta t}).\end{aligned}$$

Insgesamt:

$$q(\Delta t) = \frac{\sigma\sqrt{\Delta t} + (r - \tfrac{1}{2}\sigma^2)\Delta t + o(\Delta t)}{2\sigma\sqrt{\Delta t} + o(\sqrt{\Delta t})} = \tfrac{1}{2} + \frac{r - \tfrac{1}{2}\sigma^2}{2\sigma} \sqrt{\Delta t} + o(\sqrt{\Delta t}).$$

Herleitung der geometrischen Brownschen Bewegung unter Q An dieser Stelle ist wichtig, dass die Anzahl der Schritte N_u in denen die Aktie steigt, binomialverteilt ist. Und zwar zur Anzahl der Schritte $n = T/\Delta t$ mit festen Wahrscheinlichkeit $q(\Delta t)$. Dann gilt

$$S_T^{(n)} = S_0 u^{N_u} d^{n-N_u} = S_0 \exp(\sigma\sqrt{\Delta t} (2N_u - n)).$$

Und somit

$$\log S_T^{(n)} = \log S_0 + \sigma\sqrt{\Delta t} (2N_u - n).$$

Es wird eine 0 addiert:

$$2N_u - n = 2(N_u - nq(\Delta t)) + n(2q(\Delta t) - 1).$$

Zwischenfazit:

$$\log S_T^{(n)} = \log S_0 + \sigma\sqrt{\Delta t} n(2q(\Delta t) - 1) + \sigma\sqrt{\Delta t} 2(N_u - nq(\Delta t)).$$

Zuerst wird der stochastische Term (rechts) untersucht: Da die Binomialverteilung eine Summe von unabhängigen Bernoulli-Verteilungen ist, kann der zentrale Grenzwertsatz angewendet werden:

$$\lim_{n \rightarrow \infty} N_u = \lim_{n \rightarrow \infty} \sum_{k=0}^n \text{Ber}(q(\Delta t)) \sim \mathcal{N}(nq(\Delta t), nq(\Delta t)(1 - q(\Delta t))).$$

Das ist äquivalent zu

$$\frac{N_u - nq(\Delta t)}{\sqrt{nq(\Delta t)(1 - q(\Delta t))}} \longrightarrow Z \sim \mathcal{N}(0, 1),$$

also

$$2(N_u - nq(\Delta t)) = 2\sqrt{nq(\Delta t)(1 - q(\Delta t))} Z_n,$$

wobei $Z_n \longrightarrow Z$. Schließlich folgt aus $q(\Delta t) \rightarrow \frac{1}{2}$ für $\Delta t \rightarrow 0$:

$$\begin{aligned}\sigma\sqrt{\Delta t} \cdot 2\sqrt{nq(\Delta t)(1 - q(\Delta t))} Z_n &= 2\sigma\sqrt{n\Delta t} \sqrt{q(\Delta t)(1 - q(\Delta t))} Z_n \\ &= 2\sigma\sqrt{T} \sqrt{\tfrac{1}{4} + o(1)} Z_n = \sigma\sqrt{T} Z_n + o(1).\end{aligned}$$

Nun wird der deterministische Term (mittig) untersucht:

Multipliziere mit $\sigma\sqrt{\Delta t}$ und benutze $n\Delta t = T$ sowie $q(\Delta t) = \frac{1}{2} + \alpha\sqrt{\Delta t} + o(\sqrt{\Delta t})$

mit $\alpha = \frac{r - \frac{1}{2}\sigma^2}{2\sigma}$:

$$\begin{aligned}\sigma\sqrt{\Delta t} n(2q(\Delta t) - 1) &= \sigma\sqrt{\Delta t} n(2\alpha\sqrt{\Delta t} + o(\sqrt{\Delta t})) = \sigma n(2\alpha\Delta t + o(\Delta t)) \\ &= \sigma \cdot 2\alpha \cdot (n\Delta t) + o(1) = (r - \tfrac{1}{2}\sigma^2)T + o(1).\end{aligned}$$

Zwischenfazit:

$$\log S_T^{(n)} = \log S_0 + (r - \tfrac{1}{2}\sigma^2)T + \sigma\sqrt{T} Z_n + o(1).$$

Im Grenzwert $n \rightarrow \infty$ (bzw. $\Delta t \rightarrow 0$) folgt damit

$$\log S_T \sim \mathcal{N}\left(\log S_0 + (r - \tfrac{1}{2}\sigma^2)T, \sigma^2 T\right),$$

d. h. unter Q konvergiert der diskrete Prozess gegen die geometrische Brownsche Bewegung. Der Drift-Term ist nicht mehr der Trend im Aktienkurs sondern der Zinssatz r . \square

Satz 7.3.1 (Satz von Pratt, Elstrodt [6] S. 280, vereinfacht). Dieser technische Satz erlaubt es, den Grenzübergang und den Erwartungswert zu vertauschen, und kommt im Folgenden bei der Herleitung der Black-Scholes-Formel zum Einsatz. Sei $(X_n) \rightarrow X$ eine Folge von Zufallsvariablen, die fast-überall konvergiert, und $(Y_n) \rightarrow Y$, $(Z_n) \rightarrow Z$ ebenfalls. Gilt (1) $Y_n \leq X_n \leq Z_n$ und (2) $E(Y_n) \rightarrow E(Y)$, $E(Z_n) \rightarrow E(Z)$, dann folgt $E(X_n) \rightarrow E(X)$. Wird hier nicht bewiesen. \square

Satz 7.3.2 (Black–Scholes–Formel für europäische Call-Optionen). Sei $C_0^{(n)}$ der Preis der europäischen Call-Option mit Laufzeit T und Strike K im n -stufigen Binomialmodell unter Diskontierung mit r . Dann gilt

$$\lim_{n \rightarrow \infty} C_0^{(n)} = C_0^{\text{BS}} = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2),$$

wobei

$$d_1 = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T},$$

und Φ die Verteilungsfunktion der Standardnormalverteilung ist.

*Beweis*³. Risikoneutrale Bewertung im Binomialmodell liefert

$$C_0^{(n)} = e^{-rT} E_Q[(S_T^{(n)} - K)^+].$$

Die obige Verteilungskonvergenz impliziert

$$S_T^{(n)} \rightarrow S_T = S_0 \exp((r - \tfrac{1}{2}\sigma^2)T + \sigma\sqrt{T} Z).$$

³Der Beweis ist angelehnt an die Argumentation in Hull, 2022 [10], S. 347

Nun wird der Satz von Pratt angewendet, um den Grenzübergang und den Erwartungswert zu vertauschen. Eine untere Schranke ist $Y_n = 0$ und eine obere Schranke ist $Z_n = (S_T^{(n)})^+ = S_T^{(n)}$.

$$\lim_{n \rightarrow \infty} C_0^{(n)} = e^{-rT} E_Q \left[\left(\lim_{n \rightarrow \infty} S_T^{(n)} - K \right)^+ \right] = e^{-rT} E_Q \left[(S_T - K)^+ \right].$$

Nun wird die rechte Seite explizit berechnet: Seien

$$m := \left(r - \frac{1}{2}\sigma^2\right)T, \quad v := \sigma\sqrt{T}, \quad Z \sim \mathcal{N}(0, 1),$$

also $S_T = S_0 e^{m+vZ}$. Dann gilt

$$E_Q \left[(S_T - K)^+ \right] = E_Q \left[S_T \mathbf{1}_{\{S_T \geq K\}} \right] - K Q(S_T \geq K).$$

Aus Kapitel 5 ist bekannt, dass S_T lognormalverteilt ist. Die Schranke $S_T \geq K$ ist daher äquivalent zu $m + vZ \geq \ln(K/S_0)$, also

$$Z \geq a := \frac{\ln(K/S_0) - m}{v}.$$

Damit ergibt sich für den zweiten Term

$$Q(S_T \geq K) = Q(Z \geq a) = \Phi(-a).$$

Für den ersten Term wird die Dichte φ von Z genutzt:

$$\begin{aligned} E_Q \left[S_T \mathbf{1}_{\{Z \geq a\}} \right] &= S_0 e^m \int_a^\infty e^{vx} \varphi(x) dx = S_0 e^m \int_a^\infty e^{vx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= S_0 e^m \int_a^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x^2 - 2vx)\right) dx \\ &= S_0 e^m \int_a^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-v)^2 + \frac{1}{2}v^2\right) dx \\ &= S_0 e^{m+\frac{1}{2}v^2} \int_a^\infty \varphi(x-v) dx. \end{aligned}$$

Mit der Substitution $y = x - v$ folgt

$$\int_a^\infty \varphi(x-v) dx = \int_{a-v}^\infty \varphi(y) dy = \Phi(v-a).$$

Der zweite Term ist damit bestimmt:

$$E_Q \left[S_T \mathbf{1}_{\{S_T \geq K\}} \right] = S_0 e^{m+\frac{1}{2}v^2} \Phi(v-a).$$

Zwischenfazit:

$$E_Q \left[(S_T - K)^+ \right] = S_0 e^{rT} \Phi(d_1) - K \Phi(d_2)$$

Setze $m + \frac{1}{2}v^2 = rT$ und beachte

$$v - a = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_1, \quad -a = \frac{\ln(S_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} = d_2.$$

Damit gilt

$$E_Q[(S_T - K)^+] = S_0 e^{rT} \Phi(d_1) - K \Phi(d_2),$$

und folglich

$$C_0^{\text{BS}} = e^{-rT} E_Q[(S_T - K)^+] = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2).$$

□

Bemerkung 7.3.3. Für den europäischen Put folgt analog

$$P_0^{\text{BS}} = K e^{-rT} \Phi(-d_2) - S_0 \Phi(-d_1),$$

oder per Put-Call-Parität.

Beispiel 7.3.4 (Call-Option auf den DAX). Mit der zuvor geschätzten Volatilität (σ) des DAX kann man nun den Preis einer Call-Option auf den DAX berechnen. Das folgende R-Programm (Ausschnitt) berechnet den Preis einer Call-Option abhängig vom Ausübungspreis K .

```
1 d1 <- (log(S0/K) + (r + 0.5 * sigma^2) * T) / (sigma * sqrt(T))
2 d2 <- d1 - sigma * sqrt(T)
3 C <- S0 * pnorm(d1) - K * exp(-r*T) * pnorm(d2)
```

Ein Beispiel mit $r = 0.02$, $T = 1$ Jahr und dem DAX-Stand $S_0 = 24290$ wird in der folgenden Grafik dargestellt. Der faire Preis der Option fällt (erwartungsgemäß) mit dem Ausübungspreis K . Aus der Call-Put-Parität folgt, dass der Preis einer Put-Option sich durch

```
1 P <- C - S0 + K * exp(-r * T)
```

berechnet.

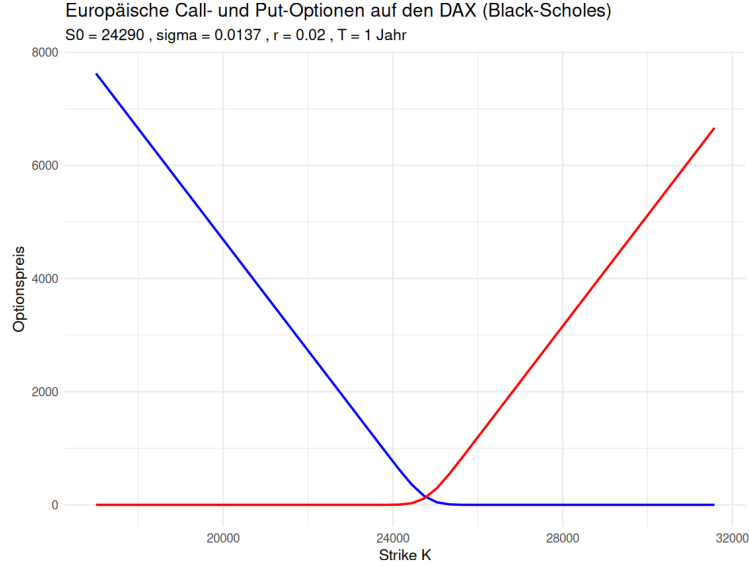


Abbildung 9: Preis einer Call-Option auf den DAX in Abhängigkeit vom Ausübungspreis K

7.4 Bewertung von Aktienoptionen mit Monte-Carlo-Verfahren

Der letzte Beweis zeigt, dass explizite Formeln für Optionspreise nur mit hohem Aufwand hergeleitet werden können. Man bedenke, dass es sich um eine vergleichsweise simple Option gehandelt hat. Für komplexere Derivate gibt es meist keine geschlossenen Formeln. Stattdessen werden numerische Verfahren verwendet, um den Erwartungswert der diskontierten Auszahlung zu approximieren.

Definition und Lemma 7.4.1 (Monte-Carlo-Verfahren). Monte-Carlo-Verfahren sind stochastische Simulationsmethoden, die zur numerischen Lösung von Problemen verwendet werden, insbesondere zur Berechnung von Integralen und Erwartungswerten. Sie basieren auf der Erzeugung von Zufallszahlen und der statistischen Analyse der Ergebnisse. Man verwendet den *Monte-Carlo-Schätzer*

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n f(X_i),$$

wobei X_1, X_2, \dots, X_n unabhängige und identisch verteilte Zufallsvariablen sind, die der Verteilung von X folgen. Der Schätzer konvergiert fast sicher gegen den wahren Erwartungswert $I = E[f(X)]$, wenn $n \rightarrow \infty$. *Beweis.* Nach dem Gesetz der großen Zahlen gilt

$$\hat{I}_n \longrightarrow E[f(X)] = I, \quad \text{fast sicher.}$$

□

Lemma 7.4.2 (Anwendung auf Optionspreise). Um allgemeine Optionen zu bewerten, muss die Auszahlungsfunktion f nicht bloß reelle Zahlen auf reelle Zahlen abbilden, sondern des gesamten Pfad S_t , $t \in [0, T]$ auf reelle Zahlen abbilden. Bei z. B. asiatischen Optionen ist der Durchschnittskurs über die Laufzeit relevant. Sei also im Folgenden

$$f : C([0, T]) \rightarrow \mathbb{R}, \tilde{S}_T \mapsto f(\tilde{S}_T)$$

eine Auszahlungsfunktion. Der faire Option mit Auszahlung $f(S_T)$ unter dem risikoneutralen Maß Q ist

$$C_0 = e^{-rT} E_Q[f(S_T)].$$

Der Monte-Carlo-Schätzer für C_0 ist somit

$$\hat{C}_0^{(n)} = e^{-rT} \frac{1}{n} \sum_{i=1}^n f(S_T^{(i)}),$$

wobei $S_T^{(i)}$, $i = 1, \dots, n$ unabhängige Realisierungen des Prozesses S_T sind. Aus Kapitel 6 ist ein Algorithmus bekannt, um Realisierungen der geometrischen Brownschen Bewegung S_T zu erzeugen.

Beispiel 7.4.3 (Monte-Carlo-Simulation einer Call-Option auf den DAX).

Das folgende R-Programm simuliert den Preis einer Call-Option auf den DAX mit dem Monte-Carlo-Verfahren. Dabei wird die geometrische Brownsche Bewegung simuliert und der Erwartungswert der diskontierten Auszahlung approximiert.

```

1  Z <- rnorm(n)
2  ST <- S0 * exp((r - 0.5*sigma^2)*T + sigma*sqrt(T)*Z)
3  payoff <- pmax(ST - K, 0) # Bewertungsfunktion
4  C <- exp(-r*T) * mean(payoff)
```

Nun wird der simulierte Preis mit dem Preis aus der Black-Scholes-Formel verglichen. Die folgende Grafik zeigt den Vergleich für verschiedene Ausübungspreise K . Je höher die Anzahl der Simulationen n , desto genauer ist die Approximation.

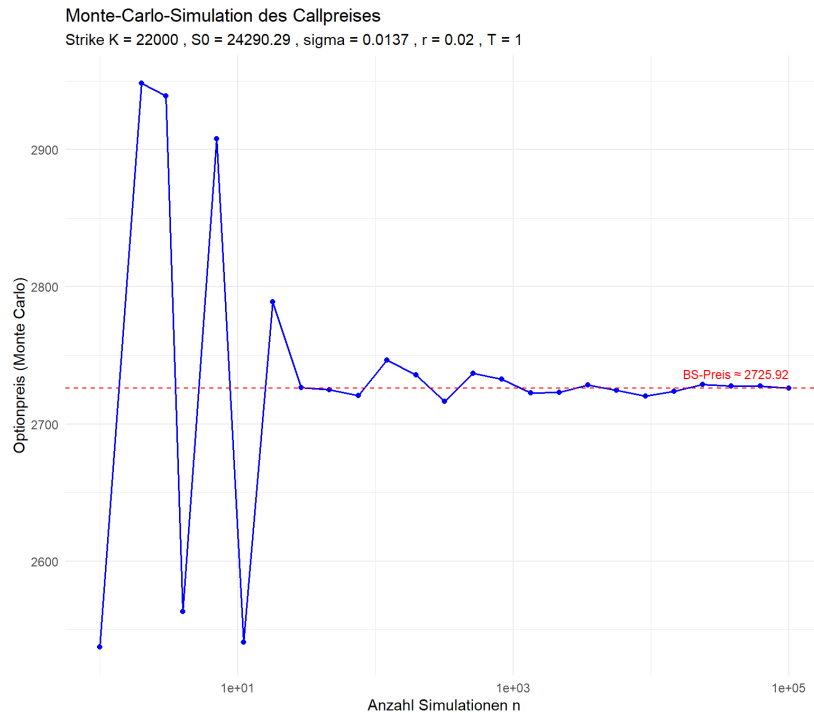


Abbildung 10: Preis einer Call-Option auf den DAX in Abhängigkeit der Simulationsanzahl n im Vergleich zur Black-Scholes-Formel

Beispiel 7.4.4 (Asiatische Optionen). Asiatische Optionen sind Finanzderivate, bei denen die Auszahlung auf dem Durchschnittspreis des Basiswerts über einen bestimmten Zeitraum basiert, anstatt auf dem Preis zu einem einzigen Zeitpunkt. Die Auszahlung einer asiatischen Call-Option mit Ausübungspreis K und Laufzeit T ist

$$f(S_T) = \max \left(\frac{1}{T} \int_0^T S_t dt - K, 0 \right).$$

Die Bewertung erfolgt mit dem Monte-Carlo-Verfahren, da es keine geschlossene Formel wie bei der Black-Scholes-Formel gibt. Die Implementierung in R könnte wie folgt aussehen:

```

1 Z <- matrix(rnorm(N * m), nrow=N, ncol=m) # N Pfade mit m
  Zeitschritten
2 dt <- T / m
3 S <- matrix(0, nrow=N, ncol=m)
4 S[,1] <- S0
5 for (j in 2:m) {
6   S[,j] <- S[,j-1] * exp((r - 0.5*sigma^2)*dt + sigma*sqrt(dt)*Z[,j])
7 }

```

```

8 avg_price <- rowMeans(S) # Durchschnittspreis für jeden Pfad
9 payoff <- pmax(avg_price - K, 0) # Bewertungsfunktion
10 C <- exp(-r*T) * mean(payoff) # diskontierter Monte-Carlo-Schätzer

```

8 Alternative Modelle

Dieses Kapitel bietet einen Einblick in weiterführende mathematische Aspekte, die im Rahmen dieser Arbeit nicht im Detail behandelt werden konnten. Die Grundlage bilden stochastische Differentialgleichungen (SDEs). Stochastische Differentialgleichungen bilden die mathematische Grundlage für die Modellierung und Simulation zeitstetiger stochastischer Prozesse im allgemeinen. Im Folgenden wird die Herleitung von Itô-SDEs über Itô-Integrale skizziert. Unter Verwendung des Kalküls der SDEs werden alternative Modelle zur geometrischen Brownschen Bewegung vorgestellt. Beispielhaft wird das CEV-Modell (Constant Elastic Variance) untersucht, implementiert und mit der GBM verglichen.

Bemerkung 8.0.1 (Motivation Stochastischer Differentialgleichungen). Im Kapitel 5 wurde ein Aktienkurs durch die zeitdiskrete Übergangsgleichung

$$S_{n+1} = S_n (1 + \mu \Delta t + \sigma \sqrt{\Delta t} \varepsilon_{n+1})$$

modelliert. Ziel des folgenden Abschnitts ist es, den Grenzübergang $\Delta t \rightarrow 0$ zu betrachten, um eine kontinuierliche Beschreibung der Kursdynamik zu erhalten. Dabei soll die stochastische Differentialgleichung

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

hergeleitet werden, deren Lösung die geometrische Brownsche Bewegung ist. Anschaulich muss man sich nur noch von $\sqrt{\Delta t} \varepsilon_{n+1} \rightarrow dW_t$ überzeugen, wobei W_t eine Brownsche Bewegung ist. Sei $\Delta t = T/n$ und $(\varepsilon_k)_{k \geq 1}$ i.i.d. mit $E(\varepsilon_k) = 0$, $V(\varepsilon_k) = 1$. Definiere den skalierten Zufallsspaziergang

$$W^{(n)}(t) := \sqrt{\Delta t} \sum_{k=1}^{\lfloor t/\Delta t \rfloor} \varepsilon_k, \quad t \in [0, T].$$

Dann gilt für jedes feste t nach dem Zentralen Grenzwertsatz

$$W^{(n)}(t) \rightarrow \mathcal{N}(0, t).$$

Und nach Kapitel 4 konvergiert $W^{(n)}$ sogar gegen eine Brownsche Bewegung W .

Insbesondere sind die diskreten Inkremente unabhängig und es gilt für $k = \lfloor t/\Delta t \rfloor$

$$W^{(n)}(t + \Delta t) - W^{(n)}(t) = \sqrt{\Delta t} \varepsilon_{k+1} \sim \mathcal{N}(0, \Delta t).$$

Mit $W^{(n)} \rightarrow W$ folgt damit für jedes t die Verteilungskonvergenz der Inkremente

$$\sqrt{\Delta t} \varepsilon_{k+1} = W^{(n)}(t + \Delta t) - W^{(n)}(t) \rightarrow W_{t+\Delta t} - W_t,$$

Durch Missbrauch der Notation erhält man

$$\sqrt{\Delta t} \varepsilon_{n+1} \rightarrow dW_t.$$

Insgesamt gilt

$$\Delta t \rightarrow dt, \quad S_{t+\Delta t} - S_t \rightarrow dS_t$$

und $\sqrt{\Delta t} \varepsilon_{n+1} \rightarrow dW_t$ kann die Übergangsgleichung als stochastische Differentialgleichung interpretiert werden. \square

Formal werden stochastische Differentialgleichungen über das Itô-Integral eingeführt, das wird im nächsten Abschnitt skizziert, für eine genaue Darstellung sei auf die Literatur verwiesen, z. B. Behrends (2013) [2] oder Karatzas und Shreve (1991) [13]. Die Argumentation ist analog zu Kapitel 6 des Lehrbuchs von Behrends [2].

8.1 Stochastische Differentialgleichungen

Ziel dieses Abschnitts ist es, die heuristische Schreibweise

$$dS_t = a(S_t, t) dt + b(S_t, t) dW_t$$

zu präzisieren. Genauso wie man eine klassische Differentialgleichung, z. B

$$dy = f(y)$$

durch die Integralgleichung

$$y = \int f(y)$$

beschreiben kann, kann man Itô-SDEs mit stochastischen Integralen (Itô-Integralen) beschreiben. Arbeitsraum ist ein filtrierter Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}),$$

und es sei W_t eine Brownsche Bewegung.

Analog zu analytischen Integralen wird das stochastische Integral über Treppenfunktionen definiert. Genauso wie man eine integrierbare Funktion mit einer Folge von Stufen approximiert, werden stochastische Prozesse durch elementare Prozesse approximiert.

Definition 8.1.1 (Elementare (adaptierte) Prozesse und ihr Integral).

Fixiere $T > 0$. Ein Prozess H heißt *elementar adaptiert* auf $[0, T]$, wenn es eine Zerlegung $0 = t_0 < t_1 < \dots < t_n = T$ und Zufallsvariablen $\xi_i \in L^2(\Omega, \mathcal{F}_{t_i}, \mathbb{P})$, $i = 0, \dots, n-1$, gibt mit

$$H_t = \sum_{i=0}^{n-1} \xi_i \mathbf{1}_{(t_i, t_{i+1}]}(t), \quad t \in [0, T].$$

Für solche H definiert das *Itô-Integral* gegen W durch

$$\int_0^T H_s dW_s := \sum_{i=0}^{n-1} \xi_i (W_{t_{i+1}} - W_{t_i}).$$

Zentrale Eigenschaften sind

$$E \left[\int_0^T H_s dW_s \right] = 0, \quad E \left[\left(\int_0^T H_s dW_s \right)^2 \right] = E \left[\int_0^T H_s^2 ds \right].$$

Dies ist die *Itô-Isometrie*. Mit der Itô-Isometrie kann man aus quadratintegrierbaren Prozessen schließen, dass das Itô-Integral existiert. Der entstandene Prozess ist ein Martingal.

Satz 8.1.2 (Dichtheit der elementaren Prozesse). Die Frage, wann ein stochastischer Prozess durch einen elementaren Prozess approximierbar ist wird im Folgenden beantwortet. Das wird wie folgt ausgedrückt: Die Menge der elementaren Prozesse liegt *dicht* im Raum der quadratintegrierbaren Prozesse. Dann gibt es zu jedem stochastischen Prozess einen elementaren stochastischen Prozess, deren Integrale sich um ein beliebiges ε unterscheiden.

Sei $L_{\text{pred}}^2(\Omega \times [0, T])$ der Funktionen-Raum aller vorhersagbaren Prozesse X mit der Norm

$$\|X\|_{2,T}^2 := E \left[\int_0^T |X_s|^2 ds \right] < \infty.$$

Dann ist die Klasse der elementaren vorhersagbaren Prozesse dicht in $L_{\text{pred}}^2(\Omega \times [0, T])$ bezüglich der Norm $\|\cdot\|_{2,T}$. (Siehe z. B. Karatzas und Shreve [13], S. 132f.) Die Dichtheit garantiert, dass zu einem beliebigen (adaptierten, quadratintegrierbaren) Prozess eine Folge von elementaren Prozessen existiert, die dagegen konvergiert.

Definition 8.1.3 (Itô-Integral). Das Itô-Integral ist nun für elementare Prozesse definiert, und für einen beliebigen Prozess existiert eine passende Folge. So wird dann das Itô-Integral für beliebige Prozesse definiert: Das Itô-Integral ist der Grenzwert der Integrale der approximierenden elementaren Prozesse. Offen bleibt aber die Frage nach der Wohldefiniertheit dieses Integral-Begriffs. Konvergieren zwei Folgen $G_t^{(n)}, H_t^{(n)}$ von elementaren Prozessen gegen einen Prozess X_t (in der Norm $\|\cdot\|_{2,T}^2$), so muss auch das Integral im Grenzwert übereinstimmen. Hier wird die oben genannte Itô-Isometrie genutzt.

Für einen allgemeinen vorhersagbaren Prozess $X \in L_{\text{pred}}^2(\Omega \times [0, T])$ wähle eine Folge elementarer Prozesse $H^{(n)}$ mit $\|H^{(n)} - X\|_{2,T} \rightarrow 0$. Dann definiert man

$$\int_0^T X_s dW_s := L^2\text{-}\lim_{n \rightarrow \infty} \int_0^T H_s^{(n)} dW_s,$$

wobei der Limes aufgrund der Itô-Isometrie existiert und nicht von der approximierenden Folge abhängt: Für zwei Folgen G und H gilt

$$\begin{aligned} E \left[\left(\int_0^T (H_s^{(n)} - G_s^{(n)}) dW_s \right)^2 \right] &= E \left[\int_0^T (H_s^{(n)} - G_s^{(n)})^2 ds \right] \\ &= \|H^{(n)} - G^{(n)}\|_{2,T}^2 \end{aligned}$$

Wegen der Dreiecksungleichung folgt

$$\|H^{(n)} - G^{(n)}\|_{2,T} \leq \|H^{(n)} - X\|_{2,T} + \|G^{(n)} - X\|_{2,T} \rightarrow 0.$$

Somit gilt

$$\int_0^T H_s^{(n)} dW_s - \int_0^T G_s^{(n)} dW_s \rightarrow 0 \quad \text{in } L^2(\Omega).$$

Daher ist das Itô-Integral wohldefiniert. □

Für jedes $t \in [0, T]$ definiert man analog den Prozess

$$\left(\int_0^t X_s dW_s \right)_{t \in [0, T]},$$

und es gilt weiterhin die Itô-Isometrie

$$E \left[\left(\int_0^T X_s dW_s \right)^2 \right] = E \left[\int_0^T X_s^2 ds \right],$$

sowie die Martingaleigenschaft von $M_t := \int_0^t X_s dW_s$. Für lokal quadratintegrierbare X erhält man das Integral durch Lokalisierung via Stoppzeiten.

Definition 8.1.4 (Stochastische Differentialgleichungen). Mit dieser Konstruktion ist die Schreibweise

$$dS_t = a(S_t, t) dt + b(S_t, t) dW_t$$

präzise zu lesen als Integralgleichung

$$S_t = S_0 + \int_0^t a(S_s, s) ds + \int_0^t b(S_s, s) dW_s,$$

wobei $b(S, \cdot)$ adaptiert und quadratintegrierbar sein muss.

8.2 Charakterisierung alternativer Kursmodelle durch stochastische Differentialgleichungen

Die GBM nimmt konstante Volatilität und lognormale Renditen an; sie erzeugt weder Sprünge noch Volatilitäts-Clustering oder kann Konjunkturperioden Modellieren. Glasserman [7] beschreibt verschiedene Ansätze zur Modellierung dieser Phänomene:

Beispiel 8.2.1 (Lokale Volatilität). Deterministisch zeit- und zustandsabhängige Volatilität:

$$dS_t = \mu(t, S_t) S_t dt + \sigma_{\text{loc}}(t, S_t) S_t dW_t.$$

Spezialfall CEV:

$$dS_t = \mu S_t dt + \sigma S_t^\beta dW_t, \quad \beta \in \mathbb{R},$$

wodurch die Volatilität bei kleinen/größeren Preisen relativ ansteigt/abnimmt.

Beispiel 8.2.2 (Stochastische Volatilität). Volatilität ist selbst ein Zufallsprozess.

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t dW_t^{(1)}, \\ dV_t &= \kappa(\theta - V_t) dt + \xi \sqrt{V_t} dW_t^{(2)}, \quad d\langle W^{(1)}, W^{(2)} \rangle_t = \rho dt, \end{aligned}$$

mit Feller-Bedingung $2\kappa\theta \geq \xi^2$ für Positivität von V_t .

Beispiel 8.2.3 (Sprung-Diffusions-Modelle). Diffusion mit seltenen Sprüngen modelliert durch Poisson-Prozesse.

$$\frac{dS_t}{S_t} = (\mu - \lambda \kappa_J) dt + \sigma dW_t + (J_t - 1) dN_t,$$

wobei N_t Poisson mit Intensität λ ist, J_t die Sprunggröße (z. B. lognormal) und $\kappa_J = E(J_t - 1)$; der Driftterm kompensiert die Sprünge.

Beispiel 8.2.4 (Regimewechsel-Modelle). Parameter schalten gemäß einer (verborgenen) Markov-Kette X_t :

$$dS_t = \mu_{X_t} S_t dt + \sigma_{X_t} S_t dW_t.$$

Erfasst Phasen wie Krisen und ruhige Märkte.

8.3 Simulation stochastischer Differentialgleichungen

Zu einer SDE

$$dX_t = a(X_t, t) dt + b(X_t, t) dW_t, \quad t \in [0, T],$$

wird ein Zeitschrittgitter $t_n = n\Delta t$ eingeführt und der kontinuierliche Prozess durch eine zeitdiskrete Approximation ersetzt. Erwartungswerte werden via Monte-Carlo über viele simulierte Pfade geschätzt.

Die Differentialgleichung (DGL) wird mit Euler (deterministisch) und Euler–Maruyama (stochastisch) diskretisiert (vgl. Bärwolff [1] u. a. S. 269, 466f.). Für die DGL $x'(t) = a(x(t), t)$ liefert das explizite Euler-Schema

$$x_{n+1} = x_n + a(x_n, t_n) \Delta t.$$

Ersetzt man nun formal dW_t durch $\Delta W_n := W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, \Delta t)$, erhält man die natürliche stochastische Erweiterung, das Euler–Maruyama-Schema:

$$\begin{aligned} X_{n+1} &= X_n + a(X_n, t_n) \Delta t + b(X_n, t_n) \Delta W_n \\ &= X_n + a(X_n, t_n) \Delta t + b(X_n, t_n) \sqrt{\Delta t} \varepsilon_{n+1}, \end{aligned}$$

mit i.i.d. $\varepsilon_n \sim \mathcal{N}(0, 1)$. Ähnlich wurde im Hauptteil die geometrische Brownsche Bewegung hergeleitet.

Beispiel 8.3.1 (Implementierung). Es folgt ein R-Programm zur Simulation eines CEV-Modells (s. u.).

```

1 simulate_cev_paths <- function(S0, mu, sigma, beta, dt, nsteps,
  npaths) {
2   S <- matrix(NA, nrow = nsteps + 1, ncol = npaths)
3   S[1, ] <- S0
4   for (i in 1:nsteps) {
5     Z <- rnorm(npaths)
6     S[i+1, ] <- S[i, ] + mu * S[i, ] * dt + sigma * (S[i, ]^beta) *
      sqrt(dt) * Z
7     S[i+1, ][S[i+1, ] <= 0] <- 1e-8
8   }
9   return(S)
10 }
```

8.4 Implementierung eines CEV (Constant Elasticity of Variance) Modell

Das CEV-Modell ist ein Spezialfall der lokalen Volatilität mit

$$dS_t = \mu S_t dt + \sigma S_t^\beta dW_t, \quad \beta \in \mathbb{R}.$$

- $\beta = 1$ ergibt GBM.
- $\beta < 1$ erhöht die relative Volatilität bei kleinen Preisen.
- $\beta > 1$ verstärkt Schwankungen bei hohen Preisen.

Mit diskreten Beobachtungen S_{t_i} im Abstand Δt liefert das Euler-Schema

$$\Delta S_i \approx \mu S_{t_i} \Delta t + \sigma S_{t_i}^\beta \sqrt{\Delta t} \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, 1).$$

Die nächste Herausforderung ist die Kalibrierung (Parameterschätzung) aus gegebenen Daten. Die Arbeit orientiert sich im folgenden an Iacus (2008, S. 122 ff.) [11]. Eine eigene Implementierung in R wird danach mit dem Softwarepaket Sim.DiffProc [8] (2020) verglichen, das man auf den selben Ansatz konfigurieren kann. Der Ansatz ist die numerische Maximierung einer Log-Likelihood-Funktion, die im Folgenden für die diskrete Übergangswahrscheinlichkeit hergeleitet wird. Das wird durch die Approximation der Übergangsdichte (eine bedingte Dichte) mittels des Euler-Maruyama-Schemas erreicht. Da sich für $\beta = 1$

eine klassische geometrische Brownsche Bewegung ergibt, können die bekannten Schätzer für Drift und Volatilität als Spezialfall betrachtet werden. Insbesondere stellen die Parameter der GBM sinnvolle Startwerte für die numerische Optimierung dar.

Lemma 8.4.1 ((Quasi) Maximum-Likelihood-Schätzung (MLE, QMLE)). Ziel ist die Parameterschätzung aus diskreten Beobachtungen S_{t_0}, \dots, S_{t_n} eines SDE $dS_t = a(S_t, t) dt + b(S_t, t) dW_t$. Grundidee: Wähle Parameter θ so, dass die beobachteten Daten unter dem Modell am wahrscheinlichsten sind. Für diskrete Beobachtungen S_{t_0}, \dots, S_{t_n} eines (Markov-)Modells gilt

$$L(\theta) = \prod_{i=0}^{n-1} p_\theta(S_{t_{i+1}}, \Delta t | S_{t_i}), \quad \ell(\theta) = \log L(\theta) = \sum_{i=0}^{n-1} \log p_\theta(S_{t_{i+1}}, \Delta t | S_{t_i}),$$

wobei $p_\theta(\cdot | \cdot)$ die bedingte Übergangsdichte in Schrittweite Δt ist. Der MLE ist

$$\hat{\theta} = \arg \max_{\theta} \ell(\theta).$$

Da p_θ (hier) nicht explizit bekannt ist, wird eine Quasi-Likelihood (QMLE)-Methode verwendet. Hier mit Hilfe der Euler-Maruyama-Approximation.

$$S_{t_{i+1}} - S_{t_i} | S_{t_i} \approx \mathcal{N}(a_\theta(S_{t_i}, t_i) \Delta t, b_\theta^2(S_{t_i}, t_i) \Delta t),$$

Dann wird die (Pseudo-)Log-Likelihood aufgestellt, und numerisch maximiert.

Satz 8.4.2 (Quasi-MLE für das CEV Modell (Euler-Pseudo-Likelihood)). Das CEV-Modell lautet

$$dS_t = \mu S_t dt + \sigma S_t^\beta dW_t.$$

Für kleine Zeitschritte Δt wird die Differentialgleichung diskretisiert:

$$S_{t_{i+1}} \approx S_{t_i} + \mu S_{t_i} \Delta t + \sigma S_{t_i}^\beta \sqrt{\Delta t} \varepsilon_{i+1},$$

wobei $\varepsilon_{i+1} \sim \mathcal{N}(0, 1)$ unabhängig sind. Definiere die Inkremente:

$$\Delta S_i := S_{t_{i+1}} - S_{t_i}$$

Nach obiger Diskretisierung gilt näherungsweise (in die man die deterministischen Terme in die Verteilung von ε reinzieht):

$$\Delta S_i | S_{t_i} \sim \mathcal{N}(\mu S_{t_i} \Delta t, \sigma^2 S_{t_i}^{2\beta} \Delta t)$$

Das heißt: Für gegebenen Kurs S_{t_i} ist der nächste Schritt normalverteilt mit Mittelwert $\mu S_{t_i} \Delta t$ und Varianz $\sigma^2 S_{t_i}^{2\beta} \Delta t$. Die Dichte einer Normalverteilung mit Mittelwert m und Varianz v an der Stelle x ist:

$$p(x) = \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{(x - m)^2}{2v}\right)$$

Setze $m = \mu S_{t_i} \Delta t$, $v = \sigma^2 S_{t_i}^{2\beta} \Delta t$, $x = \Delta S_i$. Die gemeinsame Likelihood für alle Schritte ist das Produkt der Einzeldichten:

$$L(\mu, \sigma, \beta) = \prod_{i=0}^{n-1} p(\Delta S_i | S_{t_i})$$

Da Produkte numerisch unhandlich sind, nimmt man den Logarithmus:

$$\ell(\mu, \sigma, \beta) = \sum_{i=0}^{n-1} \log p(\Delta S_i | S_{t_i})$$

Setzt man die Dichte p ein, ergibt sich die Likelihood

$$\begin{aligned} \ell(\mu, \sigma, \beta) &= \sum_{i=0}^{n-1} \left[-\frac{1}{2} \log(2\pi\sigma^2 S_{t_i}^{2\beta} \Delta t) - \frac{(\Delta S_i - \mu S_{t_i} \Delta t)^2}{2\sigma^2 S_{t_i}^{2\beta} \Delta t} \right] \\ &= -\frac{1}{2} \sum_{i=0}^{n-1} \left[\log(2\pi\sigma^2 S_{t_i}^{2\beta} \Delta t) + \frac{(\Delta S_i - \mu S_{t_i} \Delta t)^2}{\sigma^2 S_{t_i}^{2\beta} \Delta t} \right] \end{aligned}$$

□

Bemerkung 8.4.3 (Interpretation). Die Log-Likelihood misst, wie gut die Parameter (μ, σ, β) die beobachteten Kursänderungen erklären. Der erste Term bestraft große Varianzen, der zweite Term misst die Abweichung der beobachteten Inkremente von den erwarteten Werten. Die Schätzungen der Parameter erhält man durch Maximierung der Log-Likelihood:

$$(\hat{\mu}, \hat{\sigma}, \hat{\beta}) = \arg \max_{\mu, \sigma, \beta} \ell(\mu, \sigma, \beta)$$

Die Funktion wird mit numerischen Optimierungsverfahren (z.B. L-BFGS-B) maximiert.

Beispiel 8.4.4 (Implementierung in R mit dem Paket Sim Diffproc).

Die Parameterschätzung für das CEV-Modell kann mit dem R-Paket Sim.DiffProc durchgeführt werden. Die Funktion `fitsde` erlaubt die Schätzung der Drift- und Diffusionsparameter mittels Pseudo-Maximum-Likelihood auf Basis diskreter Zeitreihen. Das folgende R-Programm (Ausschnitt) zeigt die Anwendung auf DAX-Daten. Die Optimierung erfolgt mit dem Algorithmus L-BFGS-B und die Parametergrenzen sind gesetzt, um numerische Probleme zu vermeiden.

```

1 S_ts <- ts(dax$Price, deltat = dt_daily)
2
3 drift      <- expression(theta[1] * x)          # theta1 = mu
4 diffusion <- expression(theta[2] * x^theta[3])  # theta2 = sigma,
   theta3 = beta
5
6 fit_cev <- fitsde(
7   data      = S_ts,
```

```

8  drift      = drift,
9  diffusion  = diffusion,
10 start      = start_vals,
11 pmle       = "euler",
12 optim.method = "L-BFGS-B",
13 lower      = c(mu = -Inf, sigma = 1e-8, beta = 0.0),
14 upper      = c(mu = Inf, sigma = Inf, beta = 3.0)
15 )

```

Es folgt ein Backtest mit den geschätzten Parametern (vgl. Kapitel 7). Hier werden 1000 Pfade mit der Euler-Maruyama-Methode simuliert, und dann die empirischen 50%-Quantile genommen.

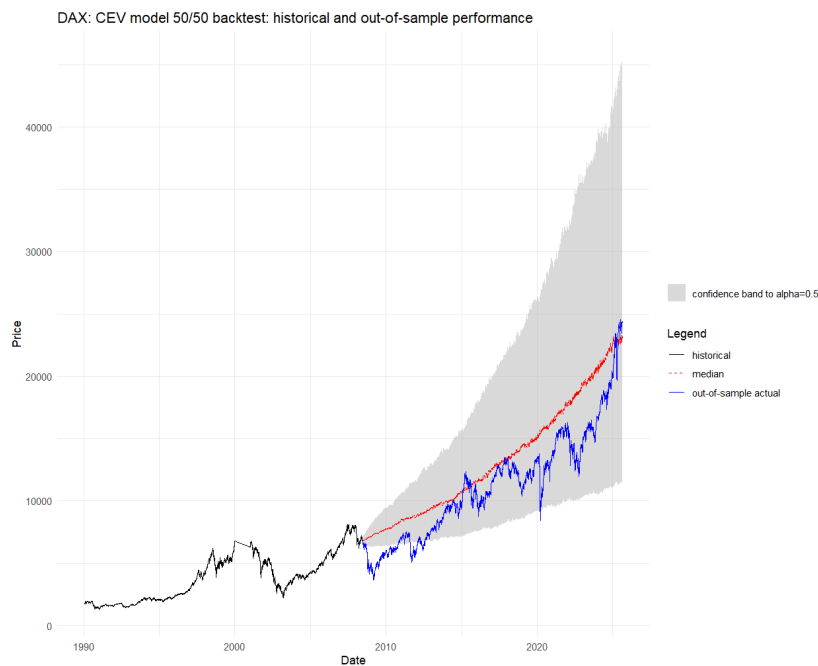


Abbildung 11: Backtest CEV-Modell für den DAX

Die Überdeckungswahrscheinlichkeit liegt bei 82%. Dieselbe Konfiguration lieferte für die geometrische Brownsche Bewegung eine Überdeckungswahrscheinlichkeit von 86% (größer ist besser). Ein feinerer Vergleich der Kalibrierungen folgt im nächsten Abschnitt.

Beispiel 8.4.5 (Eigene Implementierung in R mit der Euler-Pseudo-Likelihood).

Die eigene Implementierung der Parameterschätzung für das CEV-Modell in R basiert auf der oben hergeleiteten Pseudo-Maximum-Likelihood-Funktion. Zur numerischen Optimierung wird das Paket `nloptr` [12] verwendet.

```

1 estimate_cev <- function(S, dt = 1/252) {

```

```

2  S <- as.numeric(S)
3  stopifnot(is.numeric(S), all(is.finite(S)), length(S) >= 3)
4
5  dS <- diff(S)
6  S0 <- S[-length(S)]
7  eps <- 1e-12
8  S0[S0 <= 0] <- eps
9  sigma_start <- sd(dax_in$logret) / sqrt(dt_daily)
10 mu_start <- mean(dax_in$logret) / dt_daily - sigma_start / 2
11 beta_start <- 1.0
12 init <- c(mu_start, sigma_start, beta_start)
13
14 negloglik <- function(par) {
15   mu <- par[1]
16   sigma <- par[2]
17   beta <- par[3]
18   if (!is.finite(mu) || !is.finite(sigma) || !is.finite(beta))
19     return(1e12)
20   if (sigma <= 0 || beta <= 0 || beta >= 5) return(1e12)
21
22   denom <- (sigma*sigma) * (S0^(2*beta)) * dt
23   denom[denom <= 0] <- eps
24   val <- 0.5 * sum(log(2*pi*denom) + ((dS - mu*S0*dt)^2)/denom)
25   if (!is.finite(val)) val <- 1e12
26   as.numeric(val)
27 }
28
29 res <- nloptr(
30   x0 = init,
31   eval_f = negloglik,
32   lb = c(-Inf, 1e-8, 1e-6),
33   ub = c( Inf, Inf, 4.999),
34   opts = list(algorithm = "NLOPT_LN_SBPLX", xtol_rel = 1e-8,
35               maxeval = 3000)
36 )
37
38 list(mu = res$solution[1], sigma = res$solution[2], beta =
39       res$solution[3])

```

Vergleich der Kalibrierungen Im Vergleich des Standardpakets mit der eigenen Implementierung zeigt sich, dass die geschätzten Parameter bloß geringe Unterschiede aufweisen. Man kann schließen, dass die Schätzung korrekt implementiert ist.

Tabelle 2: Vergleich der geschätzten Parameter (μ , σ , β) für die beiden Schätzprogramme

Kurs	Eigene Implementierung			Paket		
	μ	σ	β	μ	σ	β
DAX	0.0946	0.6545	0.8726	0.0950	0.6524	0.873
Lufthansa	0.003	0.8483	0.6423	0.0035	0.8484	0.6423
Adesso	0.2686	0.4114	1.0031	0.2686	0.4122	1.0026

8.5 Ergebnisse und Vergleich der Modelle

Nun wird das CEV-Modell (unter Verwendung der in Sim.Diffproc vorgenommenen Kalibrierung) mit der geometrischen Brownschen Bewegung (GBM) verglichen. Beide Modelle werden auf denselben Datensatz angewendet und anschließend in einem Backtest bewertet.

Zum einen erfolgt eine Aufteilung der Daten nach einem bestimmten Gewicht (weight): Das Modell wird auf Basis der historischen Daten kalibriert und anschließend zur Prognose der zukünftigen Daten eingesetzt. Die Prognosegüte wird gemessen und über alle Gewichte hinweg als Mittelwert angegeben. Zum anderen wird ein sequentieller Backtest mit einem festen Zeitintervall von 20 Tagen durchgeführt, bei dem jeweils alle bis dahin verfügbaren Daten zur Prognose herangezogen werden.

Interpretation der Ergebnisse Für den DAX zeigen GBM und CEV ein sehr ähnliches Verhalten. Die Prognosequalität verbessert sich erwartungsgemäß mit zunehmendem Gewicht. Dies steht im Einklang mit der Parameterschätzung, da β nahe bei 1 liegt (s. o.) und sich das CEV-Modell in diesem Fall nahezu wie eine GBM verhält. Auch bei der Lufthansa-Aktie liefern beide Modelle vergleichbare, jedoch insgesamt schlechtere Ergebnisse, da der Kurs (anders als beim DAX) keinen (log-)linearen Trend aufweist. Als Beispiel für eine ausgeprägte Korrelation zwischen Preis und Volatilität, bei der das CEV-Modell Vorteile bietet, wird der Wechselkurs der türkischen Lira gegenüber dem US-Dollar während der Finanzkrise 2021 untersucht (vgl. Abbildung 12). Hier zeigt das CEV-Modell eine deutliche Verbesserung.

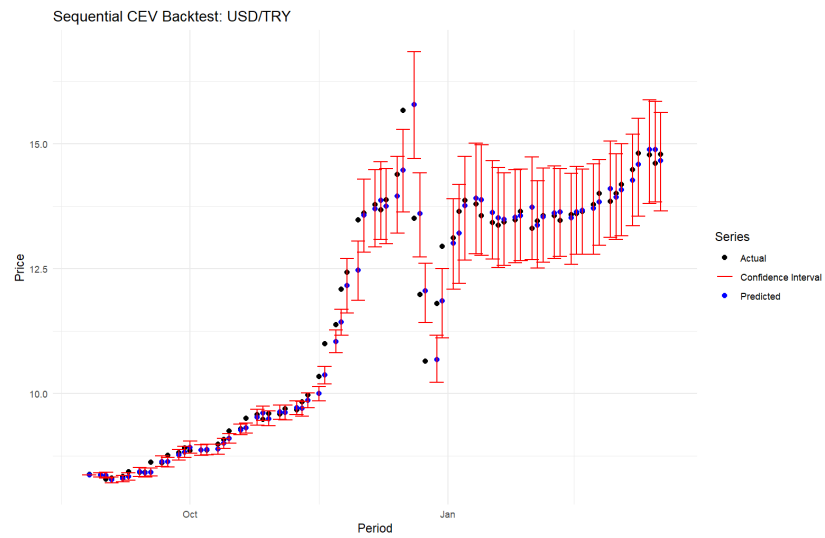


Abbildung 12: Sequenzieller Backtest des CEV Modell für die türkische Lira (Wechselrate Dollar), Abstand von 2 Tagen während der türkischen Finanzkrise 2021

Asset	Weight	CEV	GBM	CEV	GBM	CEV	GBM	CEV	GBM
dax	0.7	HitRatio	HitRatio	RMSE	RMSE	MAPE	MAPE	NRMSE	NRMSE
dax	0.75	0.681	0.696	4553.571	4888.333	30.365	32.183	0.325	0.349
dax	0.8	0.97	0.971	2306.103	2156.366	13.618	12.694	0.157	0.147
dax	0.85	0.903	0.916	1746.311	1780.029	8.751	9.181	0.114	0.117
dax	0.9	0.837	0.857	2610.952	2686.951	10.776	11.045	0.159	0.164
dax	0.95	0.8	0.805	2257.903	2543.336	8.481	9.436	0.13	0.147
dax	Seq	0.429	0.539	3513.141	3413.421	13.538	13.06	0.176	0.171
dax	mean	0.317	0.328	446.59	444.66	4.281	4.275	0.059	0.059
lufthansa	0.7	0.77	0.797	2831.33	2911.406	14.255	14.6	0.177	0.182
lufthansa	0.75	0.679	0.495	3.126	2.602	37.716	30.604	0.437	0.363
lufthansa	0.8	0.843	0.852	2.136	1.348	23.397	12.954	0.299	0.189
lufthansa	0.85	0.809	0.781	1.827	1.471	18.906	13.852	0.25	0.201
lufthansa	0.9	0.189	0.065	3.423	3.845	47.758	53.686	0.468	0.525
lufthansa	0.95	0.886	0.748	0.983	1.125	13.383	15.431	0.147	0.168
lufthansa	Seq	0.891	0.825	0.835	0.806	8.923	8.687	0.123	0.118
lufthansa	mean	0.268	0.254	1.244	1.249	7.985	8.013	0.102	0.103
tlira	0.7	0.716	0.628	2.055	1.866	25.014	22.536	0.287	0.261
tlira	0.75	0.978	1	0.629	1.116	4.126	7.269	0.045	0.08
tlira	0.8	0.974	1	0.674	0.839	4.502	5.617	0.048	0.06
tlira	0.85	0.968	1	0.413	0.428	2.673	2.844	0.029	0.03
tlira	0.9	0.957	1	0.134	0.142	0.779	0.785	0.009	0.01
tlira	0.95	0.933	1	0.242	0.29	1.325	1.689	0.017	0.02
tlira	Seq	0.875	1	0.252	0.216	1.492	1.253	0.017	0.015
tlira	mean	0.167	0.167	1.567	1.736	8.577	10.171	0.126	0.139
tlira		0.947	1	0.391	0.505	2.483	3.242	0.028	0.036

Tabelle 3: Vergleich der Modelle GBM und CEV über verschiedene Backtests und Metriken: Hitratio - größer ist besser; RMSE - kleiner ist besser; MAPE - kleiner ist besser; NRMSE - kleiner ist besser

9 Fazit

Die Arbeit spannte den Bogen von elementaren stochastischen Prozessen über das Binomialmodell und die diskrete Brownsche Bewegung bis zur geometrischen Brownschen Bewegung als kontinuierlichem Grenzfall. Zentrale stochastische Begriffe wie Filtration, bedingter Erwartungswert und Martingal wurden eingeführt und in diskreten Wahrscheinlichkeitsräumen verankert. Durch Grenzwertbetrachtungen werden die Konzepte dann auf kontinuierliche Wahrscheinlichkeitsräume übertragen. Es wurde gezeigt, dass im Grenzübergang $\Delta t \rightarrow 0$ die Kursdynamik durch $S_T = S_0 \exp((\mu - \frac{1}{2}\sigma^2)T + \sigma W_T)$ beschrieben wird, sodass $\log S_T$ normal- und S_T log-normalverteilt ist.

Im erweiterten Binomialmodell wurde der diskontierte Aktienkurs als Martingal unter einem risikoneutralen Maß konstruiert und die Optionsbewertung via Rückwärtsinduktion hergeleitet; im Grenzfall führt dies zum Black-Scholes-Modell.

Empirisch wurden die Parameter aus Log-Renditen geschätzt, daraus Konfidenzintervalle und -bänder abgeleitet und mittels Monte-Carlo-Simulation validiert. Ein Backtest auf DAX-Daten zeigte eine hohe Überdeckungsrate im 50%-Band und nachvollziehbare Fehlermaße (MSE, MAPE, NRMSE), was die Praxisauglichkeit trotz der Modellvereinfachungen unterstreicht.

Letztlich wurde gezeigt, wie die diskrete Modellierung durch Grenzübergang zur stochastischen Differentialgleichung führt, deren Lösung die geometrische Brownsche Bewegung ist. Die Konstruktion des Itô-Integrals wird skizziert und die heuristische Schreibweise $dS_t = a(S_t, t) dt + b(S_t, t) dW_t$ mathematisch präzisiert.

Darauf aufbauend werden alternative Modelle wie lokale und stochastische Volatilität (z.B. CEV- und Heston-Modell), Sprung-Diffusions- und Regimewechselmodelle vorgestellt, die realistische Markteigenschaften wie Volatilitäts-Clustering oder Sprünge abbilden können. Am Beispiel des CEV-Modells wird die Parameterschätzung aus diskreten Daten mittels (Quasi-)Maximum-Likelihood erläutert und in R implementiert. Ein Backtest und ein Vergleich mit der geometrischen Brownschen Bewegung zeigen, dass komplexere Modelle wie das CEV-Modell bei großen Datenmengen eine bessere Prognosegüte liefern, während die GBM bei kleineren Datensätzen robuster ist. Die Ergebnisse unterstreichen die Bedeutung der Modellauswahl und Kalibrierung für die praktische Anwendung in der Finanzmathematik.

Methodik Methodisch verbindet die Arbeit diskrete Grenzwert- und Martingalargumente mit reproduzierbarer Empirie in R. Zudem spielte die Monte-Carlo-Simulation eine zentrale Rolle als universelles Werkzeug zur Approximation risikoneutraler Erwartungswerte, sowohl für Endwert-Auszahlungen als auch für pfadabhängige Produkte. Für die geometrische Brownsche Bewegung wurden Endwerte unter dem risikoneutralen Maß exakt gezogen; bei Pfadabhängigkeiten wurde zeitlich diskretisiert. Die beobachtete Annäherung der empirischen Quantile an die analytischen Lognormal-Quantile mit wachsender Pfadzahl bestätigt Konsistenz und Korrektheit der Implementierung; der Vergleich analy-

tischer Konfidenzbänder mit Simulationsquantilen dient als robuster Plausibilitätscheck. Als Beispiel dient ein Datensatz von DAX-Renditen.

Die geometrische Brownsche Bewegung wird über Logarithmierung, Taylor-Entwicklung bis Ordnung zwei sowie Gesetz der großen Zahlen und zentralen Grenzwertsatz hergeleitet. Zudem kommen die Sätze von Cramér-Wold und Pratt zum Einsatz, genauso wie das "Reihenkriterium für fast sichere Konvergenz", welches ein Korollar zum Lemma von Borel-Cantelli ist (vgl. [9]). Insbesondere wurde das Kalkül der stochastischen Differentialgleichung zunächst vermieden.

Einen Einblick in die Thematik bietet Kapitel 8: die Methodik wird um die numerische Simulation und Parameterschätzung stochastischer Differentialgleichungen (SDEs) mittels Euler-Maruyama-Verfahren und Maximum-Likelihood-Ansätzen erweitert. Die Implementierung alternativer Modelle die durch SDEs beschrieben werden wurde am Beispiel des CEV-Modell in R durchgeführt. Dabei kamen die Pakete `Sim.DiffProc` und das numerische Optimierungspaket `nloptr` zum Einsatz. Anschließend wurde das CEV-Modell mit der GBM verglichen, inklusive Backtests mit empirischen Gütemaßen.

10 Quellenverzeichnis

Literatur

- [1] Günter Bärwolff und Caren Tischendorf. *Numerik in der Physik, Ingenieurwissenschaft und Informatik*. 5. Aufl. Springer, 2025. ISBN: 978-3-662-70670-1. DOI: <https://doi.org/10.1007/978-3-662-70670-1>. (Besucht am 03.09.2025).
- [2] Ehrhard Behrens. *Markovprozesse und stochastische Differentialgleichungen: Vom Zufallsspaziergang zur Black-Scholes-Formel*. Wiesbaden: Springer, 2013. ISBN: 978-3-658-00987-8.
- [3] Alexei Botchkarev. „Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology“. In: *Interdisciplinary Journal of Information, Knowledge, and Management* 14 (2019). arXiv:1809.03006 [stat], S. 045–076. ISSN: 1555-1229, 1555-1237. DOI: 10.28945/4184. URL: <http://arxiv.org/abs/1809.03006> (besucht am 30.08.2025).
- [4] Stéphane Boucheron, Gábor Lugosi und Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. en. Oxford University Press, Feb. 2013. ISBN: 978-0-19-953525-5. DOI: 10.1093/acprof:oso/9780199535255.001.0001. URL: <https://academic.oup.com/book/26549> (besucht am 19.09.2025).
- [5] *Definition of BROWNIAN MOTION, Webster's Dictionary*. URL: <https://www.merriam-webster.com/dictionary/Brownian+motion> (besucht am 30.08.2025).

-
- [6] Jürgen Elstrodt. *Maß- und Integrationstheorie*. 8. Aufl. Springer, 2011. ISBN: 978-3-662-57939-8.
 - [7] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. 1. Aufl. Bd. 53. Stochastic Modelling and Applied Probability. Springer Science+Business Media New York 2003. Part of the Springer Book Archive. New York, NY: Springer, 2003, S. XIII+596. ISBN: 978-0-387-00451-8. DOI: 10.1007/978-0-387-21617-1.
 - [8] Arsalane Chouaib Guidoum und Kamal Boukhetala. „Performing Parallel Monte Carlo and Moment Equations Methods for Itô and Stratonovich Stochastic Differential Systems: R Package Sim.DiffProc“. In: *Journal of Statistical Software* 96.2 (2020), S. 1–82. DOI: 10.18637/jss.v096.i02. URL: <https://cran.r-project.org/web/packages/Sim.DiffProc/> (besucht am 03.09.2025).
 - [9] Norbert Henze. *Stochastik: Eine Einführung mit Grundzügen der Maßtheorie*. 2. Aufl. Springer, 2023. ISBN: 978-3-662-68649-2.
 - [10] John C. Hull. *Options, Futures, and Other Derivatives*. 11. Aufl. Pearson, 2022. ISBN: 978-0-13-693997-9.
 - [11] Stefano M. Iacus. *Simulation and Inference for Stochastic Differential Equations: With R Examples*. Springer, 2008. ISBN: 978-0-387-75838-1. DOI: 10.1007/978-0-387-75839-8. URL: <https://doi.org/10.1007/978-0-387-75839-8> (besucht am 03.09.2025).
 - [12] Steven G. Johnson. *The NLOpt nonlinear-optimization package*. 2008. URL: <https://github.com/stevengj/nlopt>.
 - [13] Ioannis Karatzas und Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. 2. Aufl. Bd. 113. Graduate Texts in Mathematics. ISSN: 0072-5285. New York, NY: Springer, 1991. ISBN: 978-0-387-97655-6. DOI: 10.1007/978-1-4612-0949-2.
 - [14] Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. 1. Aufl. Springer, 2004. ISBN: 978-1-4419-2311-0.