



DYNAMISCHE MONTE CARLO METHODEN

BACHELORARBEIT

Vorgelegt von:
Patrick Wolff

Westfälische Wilhelms-Universität Münster
Fachbereich Mathematik und Informatik

Betreuung:
PD Dr. Volkert Paulsen

Münster, 08.August 2011

Eidesstattliche Erklärung

Hiermit versichere ich, Patrick Wolff, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Sinngemäß oder wörtlich Übernommenes habe ich durch Angabe von Herkunft oder Anmerkung belegt bzw. kenntlich gemacht. Darüber hinausgehend versichere ich, dass ich angefügte Skizzen, Tabellen und Matlab-Codes selbst erstellt habe.

Münster, den 08.August 2011

Inhaltsverzeichnis

1	Einleitung	4
2	Mathematische Grundlagen	6
2.1	Einblick in Statische Monte Carlo Methoden	6
2.1.1	Erzeugung von Samples von Zufallsvariablen	7
2.1.2	Acceptance-Rejection Verfahren	8
2.2	Stochastische Analysis	10
2.2.1	Stochastische Prozesse	10
2.2.2	Das Itô-Integral	10
2.2.3	Die Itô-Formel	17
2.2.4	Spezielle SDEs und „strong solutions“	18
3	Dynamische Monte Carlo Methoden	21
3.1	Diskretisierungsverfahren	21
3.2	Simulation einer Brownschen Bewegung	23
3.3	Simulation einer geometrischen Brownschen Bewegung	24
3.4	Jump Prozesse	26
3.5	Mixed Jump Prozesse	29
4	Bewertung Asiatischer Optionen	34
4.1	„Naive“ Bewertung	35
4.2	Varianzreduktionsmethoden	36
4.2.1	Kontrollvariablen	37
4.2.2	Importance Sampling	41
4.3	Berechnung von Preisen und Fehlerabschätzung	46
4.4	Put-Call-Parität	47
4.5	Das stochastische Volatilitätsmodell nach Heston	49
5	Fazit	52
6	Anhang	53

1 Einleitung

Der erste Gedanke, welcher den meisten Menschen in den Kopf kommt, wenn sie den Begriff „Monte Carlo“ hören, ist wahrscheinlich das berühmte Spielkasino in Monte Carlo. Die mathematischen Methoden, die sich hinter „Monte Carlo“ verbergen, bekamen genau wegen des Zusammenhangs von Glücksspiel und statistischer Simulation ihren Namen. Ganz allgemein kann man sagen, dass man bei Monte Carlo Methoden zufällige Zahlen erzeugt, um mit ausreichend vielen Zufallszahlen Rückschlüsse auf Erwartungswerte ziehen zu können. In der Mathematik verbindet man mit dem Begriff „Monte Carlo“ allgemein gesagt also Methoden, um statistische Experimente zu simulieren, d.h. mögliche Ausgänge von statistischen Experimenten hervorzubringen. Normalerweise erhält man dabei mittels einer gegebenen Wahrscheinlichkeitsverteilung, welche die statistischen Eigenschaften des Experiments beschreibt, ein sogenanntes Sample, eine Probe bzw. ein Beispiel dieser Verteilung. Ist es nicht möglich oder zu aufwändig ein Experiment zu wiederholen, wird zunächst ein Modell des eigentlichen Experiments erstellt, um dann eine Simulation auf diesem Modell durchzuführen. Das erhaltene Ergebnis stellt dann eine Annäherung einer Experimentdurchführung des eigentlichen Experiments (also eine simulierte Stichprobe) dar. In dieser Arbeit werden wir dabei Szenarien über die zukünftige Entwicklung von Finanzmarktvariablen, im speziellen von „risky assets“ bzw. Aktien, generieren, indem wir Samples von Zufallsvariablen und stochastischen Prozessen in einem Marktmodell mithilfe von Zufallszahlen erzeugen.

Um dies zu präzisieren, wollen wir der Frage nachgehen, was genau simuliert werden soll und warum eine Simulation in der Finanzmathematik überhaupt nützlich sein kann. Ein Ziel der Finanzmathematik ist es auf Basis eines gewissen Modells des Finanzmarktes Preise von Claims bzw. Derivaten zu berechnen. Solche Preise zum Zeitpunkt t kann man meist als Erwartungswerte der Form

$$\mathbb{E}[e^{-rT}V(T)|\mathcal{F}_t]$$

bzgl. eines äquivalenten Martingalmaßes \mathbb{P} darstellen, wobei r die risikolose Zinsrate, T die Laufzeit des Claims und $V(T)$ eine Funktion, welche den Wert des „underlying assets“ angibt, beschreiben. Oft ist es nicht möglich, solche Erwartungswerte analytisch zu berechnen. Dies ist beispielsweise häufig der Fall, wenn man sogenannte pfadabhängige Optionen betrachtet, deren Wert in T nicht nur vom aktuellen Wert des underlying assets, sondern von vielen Werten des assets über den Betrachtungszeitraum hinweg abhängt. Genau aus diesem Grund besitzt die Monte Carlo Simulation zur Schätzung derartiger Erwartungswerte in der finanzmathematischen Anwendung eine große Bedeutung.

Der Aufbau der Arbeit gestaltet sich wie folgt. Das erste Kapitel nach der Einleitung gibt eine Einführung in mathematische Grundlagen. Wir grenzen dazu zunächst dynamische und statische Monte Carlo Verfahren voneinander ab und geben einen kurzen Einblick in die statischen Methoden. Darauf folgt eine grobe Einführung in stochastische Prozesse und Itô-Rechnung, welche wir benötigen, um

mit Begriffen wie Pfade von stochastischen Prozessen und Integralen bzgl. eines Wiener Prozesses umgehen zu können. In Kapitel 3 werden wir dynamische Monte Carlo Methoden einführen. Im anschließenden Kapitel werden wir dynamische Monte Carlo Methoden nutzen, um Asiatische Optionen im Black-Scholes Modell zu bewerten. Des weiteren werden wir in dem Kapitel Methoden vorstellen, mit denen wir die Genauigkeit der Schätzungen erhöhen können. Abschließend bewerten wir Asiatische Optionen in einem stochastischen Volatilitätsmodell, dem Heston Modell.

2 Mathematische Grundlagen

2.1 Einblick in Statische Monte Carlo Methoden

Monte Carlo Methoden stellen, wie wir oben beschrieben haben, eine Möglichkeit dar, Erwartungswerte zu schätzen. Wollen wir also im einfachsten Fall den Erwartungswert einer integrierbaren Zufallsvariable X mit gegebener Verteilung auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ schätzen, erzeugen wir $n \in \mathbb{N}$ Samples, also Zufallswerte von X , summieren diese auf und teilen durch n :

$$\mathbb{E}(X) \approx \left(\frac{1}{n} \sum_{i=1}^n X^{(i)}\right),$$

wobei die $X^{(i)}$ unabhängige identisch verteilte Samples von X sind. Das starke Gesetz der großen Zahlen liefert uns dann für $n \rightarrow \infty$ die Konvergenz von $\frac{1}{n} \sum_{i=1}^n X^{(i)}$ gegen den Erwartungswert μ von X (welcher wegen der Integrierbarkeit von X existiert).

Diese Methode ist in der Tat schon eine einfache Monte Carlo Methode und verkörpert die Grundidee dieser Verfahren.

Kommt der Begriff Monte Carlo in der Stochastik bzw. der Wahrscheinlichkeitstheorie auf, geht es meist um die Integration mithilfe von Monte Carlo. Betrachten wir also ein Integral I auf dem Einheitsintervall $[0,1]$ der Form $I := \int_0^1 f(x)dx$, dann können wir dieses auch als Erwartungswert $E[f(U)]$, wobei U eine auf dem Intervall $[0,1]$ gleichverteilte Zufallsvariable sei, schreiben. Bestimmen wir also n Samples der Gleichverteilung auf $[0,1]$ U_1, \dots, U_n , so erhalten wir eine Schätzung für den Erwartungswert, und zwar

$$I_n := \frac{1}{n} \sum_{i=1}^n f(U_i).$$

Die Konvergenz gegen den eigentlich Wert des Integrals erhalten wir dann, wie oben beschrieben, mit dem Gesetz der großen Zahlen.

Ist f zweifach integrierbar, so folgt weiter mit dem zentralen Grenzwertsatz, dass

$$\sqrt{n}(I_n - E[f(U)]) \xrightarrow{d} \mathcal{N}(0, \sigma_f^2),$$

wobei $\sigma_f^2 = \int_0^1 (f(x) - E[f(U)])^2 dx$ die Varianz von f sei. Ist σ_f nicht gegeben, so kann man σ_f genau wie $E[f(U)]$ approximieren. Wir erhalten dann eine Verteilung für den Schätzfehler

$$\epsilon := |I_n - I| \sim \frac{\sigma_f}{\sqrt{n}} \mathcal{N}(0, 1).$$

Der Fehler ist also von der Ordnung $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$. Dies bedeutet, dass, wenn wir den Fehler um den Faktor k verringern wollen, wir die Anzahl von erzeugten Samples um den Faktor k^2 erhöhen müssen. Vergleichen wir dies zum Beispiel mit der Trapezregel für die Integration, bei welcher der Fehler von der Ordnung $\mathcal{O}\left(\frac{1}{n^2}\right)$ ist, so wird uns schnell der große Nachteil der Monte Carlo Methoden ersichtlich, nämlich die im Vergleich zu anderen numerischen Verfahren langsame Konvergenz

bzw. hohe Ungenauigkeit. Warum besitzen Monte Carlo Methoden dennoch große Bedeutung im Finanzwesen?

Der Vorteil von Monte Carlo Methoden ist, dass die soeben aufgezeigte Ordnung der Konvergenz der Monte Carlo Integration für alle Dimensionen d dieselbe, also $\mathcal{O}(\frac{1}{\sqrt{n}})$, ist. Das heißt für mehrdimensionale Integrale über $[0, 1]^d$ oder sogar \mathbb{R}^d bleibt der Fehler von der gleichen Ordnung; auch wenn sich die von f abhängige Standardabweichung σ_f ändern kann. Ziehen wir nochmal den Vergleich zur Trapezregel, so kann man dort sehen, dass der Fehler bei der Berechnung von höherdimensionalen Integralen von der Ordnung $\mathcal{O}(\frac{1}{n^{2/d}})$ ist.

Wollen wir beispielsweise Asiatische oder Barriere-Optionen betrachten, so müssen wir Simulationen des kompletten Pfades des underlying assets erstellen. Die Dimension unseres Problems ist dann normalerweise mindestens so hoch wie die Anzahl der Schritte in der Simulation. Monte Carlo Verfahren, welche Pfade, also im Prinzip eine von der Zeit abhängige Realisation bzw. eine Folge von Samples von Zufallsvariablen $(X_t)_t$, simulieren, heißen dynamische Monte Carlo Verfahren. Generell kann man dabei ähnlich vorgehen wie bei der Schätzung von Zufallsvariablen mit statischen Monte Carlo Methoden. Man bestimmt zunächst ein „sehr großes“ $n \in \mathbb{N}$, nachfolgend erzeugt man n unabhängige Pfade $X_{[t,T]}^{(1)}, \dots, X_{[t,T]}^{(n)}$ des Prozesses X auf $[t, T]$. Anschließend bestimmt man für jeden erzeugten Pfad den abgezinsten Payoff des Claims oder Derivats basierend auf dem erzeugten Pfad und teilt die Summe dieser durch n . Als Ergebnis erhält man eine Schätzung des Wertes des Claims oder Derivats zum Zeitpunkt t . Deshalb wollen wir in dieser Arbeit insbesondere der Frage nachgehen, wie wir Pfade von stochastischen Prozessen simulieren können.

Darauf werden wir nach der Einführung in stochastische Analysis im Kapitel über dynamische Monte Carlo Methoden bzw. bei der Bewertung von Asiatischen Optionen weiter eingehen. Kommen wir nun erstmal zu den grundlegenden Voraussetzungen, welche wir für die Simulationen benötigen.

2.1.1 Erzeugung von Samples von Zufallsvariablen

Im Folgenden werden wir der Frage nachgehen, wie wir die Samples einer Zufallsvariablen X erzeugen können, wobei wir uns an [1] orientieren wollen. Zunächst sei bemerkt, dass es nicht möglich ist, wirklich zufällige Zahlen zu erzeugen, da Computer natürlich immer bestimmte Algorithmen benutzen müssen, um Zahlen zu erzeugen. Allerdings gibt es Zahlenfolgen, welche aussehen als wären sie wirklich zufällig erzeugte unabhängige Kopien der Gleichverteilung auf dem Einheitsintervall. Man nennt diese Zahlen deshalb auch „Pseudo-Zufallszahlen“. Ausgehend von den Pseudo-Zufallszahlen der Gleichverteilung können wir Pseudo-Zufallszahlen beliebiger anderer Verteilungen mit Hilfe von passenden deterministischen Transformationen erzeugen. Dazu zitieren wir folgendes Lemma und dessen Beweis aus [3]:

Lemma 2.1 (Pseudo-Inverse).

Sei Q eine Verteilung auf $(\mathbb{R}, \mathcal{B})$ mit Verteilungsfunktion F und

$$F^{-1}(y) := \inf\{x \in \mathbb{R} : F(x) \geq y\}, y \in (0, 1),$$

deren Pseudo-Inverse. Dann gilt $F^{-1}(U) \sim Q$ für jede auf $(0, 1)$ gleichverteilte Zufallsgröße U .

Beweis. F^{-1} erfüllt offensichtlich für alle $x \in \mathbb{R}$ die Beziehung

$$\{y \in (0, 1) : F^{-1}(y) \leq x\} = \{y \in (0, 1) : F(x) \geq y\}.$$

Daraus folgt, dass für jede auf $(0, 1)$ gleichverteilte Zufallsgröße U und für alle $x \in \mathbb{R}$ gilt

$$\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x), \text{ d.h. } F^{-1}(U) \sim Q$$

□

2.1.2 Acceptance-Rejection Verfahren

Die oben vorgestellte Methode, bei der man Samples der Gleichverteilung \mathcal{U} erstellt und diese mit Hilfe einer Pseudo-Inversen in Samples anderer Verteilungen X „umwandelt“, nennt man Inversionsmethode. Ist es schwierig, Samples von X mithilfe der Inversionsmethode zu erzeugen, etwa wenn die gegebene Verteilungsfunktion \mathbb{P}^X oder deren Pseudo-Inverse zu komplex ist, ist es oft vorteilhaft, Samples mit dem Acceptance-Rejection-Verfahren zu erzeugen.

Sei also \mathbb{P}^X die gegebene Verteilung auf dem Bildraum $\mathcal{D} := \text{Im}(X)$ von X . Wir wählen eine Zufallsvariable Y mit Verteilung \mathbb{P}^Y und Träger $\text{supp}(\mathbb{P}^Y) = \mathcal{D}$. Das heißt, alle Intervalle $I \in \text{Im}(X)$ besitzen unter \mathbb{P}^Y positive Wahrscheinlichkeit. Als nächstes unterzieht man diese Samples einem Zufallstest. Der Test ist dabei natürlich so konstruiert, dass Samples von Y , unter der Bedingung, dass sie den Test „bestehen“, die selbe Verteilung besitzen wie Samples von X . Deshalb ist dieses Verfahren insbesondere dann vorzuziehen, wenn die Erzeugung eines Samples der Verteilung \mathbb{P}^Y deutlich leichter ist als die eines Samples von X . Der offensichtliche Nachteil des Verfahrens ist ein deutlich erhöhter Zeitaufwand, der genau dann möglich ist, wenn der Test häufig negativ ausfällt.

Wir stellen das Verfahren nun im Detail vor.

Wie schon oben erwähnt, ist es das Ziel eine Verteilung \mathbb{P}^Y und eine Testfunktion $\varphi : \text{supp}(\mathbb{P}^Y) \rightarrow [0, 1]$ zu finden, so dass unter der Bedingung, dass der Zufallstest $\mathbf{1}_{\{U \leq \varphi(Y)\}}$, wobei $U \sim \mathcal{U}[0, 1]$ ein unabhängiges Sample der Gleichverteilung \mathcal{U} auf $[0, 1]$ darstellt, positiv ausfällt (also 1 ergibt), für die zugehörigen Dichtefunktionen gilt:

$$f_X(x) = f_{Y|U \leq \varphi(Y)}(x) \quad \forall x \in \mathcal{D}.$$

Satz 2.2. Seien f_X und f_Y die zu den Verteilungen von X und Y gehörigen Dichten auf \mathbb{R} und es gelte für eine Konstante $c > 0$: $f_X(x) \leq cf_Y(x) \quad \forall x \in \mathcal{D}$. Definiert man dann $\varphi(x) := f_X(x)/(cf_Y(x))$ für alle x aus \mathcal{D} , so gilt

$$f_X(x) = f_{Y|U \leq \varphi(Y)}(x) \quad \forall x \in \mathcal{D}$$

Bevor wir mit dem Beweis der Aussage beginnen, wollen wir die Aussage verdeutlichen. Der Satz sagt im Prinzip Folgendes aus: Wir müssen eine Zufallsvariable Y mit Verteilung \mathbb{P}^Y finden, welche die Bedingung $f_X(x) \leq cf_Y(x)$ für eine Konstante c und für alle $x \in \mathcal{D}$ erfüllt. Definieren wir dann φ wie im Satz, so haben wir eine Zufallsvariable und einen Zufallstest gefunden, welche die Erzeugung von Samples von X mit der Erzeugung von Samples von Y , welche man diesem Test unterzieht, gleichsetzt. Kommen wir nun zum Beweis des Satzes.

Beweis. Zunächst bemerken wir, dass $\text{supp}(\mathbb{P}^Y) = \mathcal{D} = \text{Im}(X) \supseteq \text{supp}(\mathbb{P}^X)$ und somit φ wohldefiniert ist, denn es gilt $f_Y(x) > 0$ für alle $x \in \text{supp}(\mathbb{P}^Y)$ und damit auch $cf_Y(x) > 0 \forall x \in \mathcal{D}$. Sei nun $I := [a, b] \subseteq \mathcal{D}$ ($=\text{Im}(X)$) ein beliebiges Intervall. Dann gilt:

$$\begin{aligned}
\int_a^b f_{Y|U \leq \varphi(Y)}(x) dx &= \mathbb{P}(Y \in [a, b] | U \leq \varphi(Y)) \\
&\stackrel{(*)}{=} \frac{\mathbb{P}(U \leq \varphi(Y) | Y \in [a, b]) \mathbb{P}(Y \in [a, b])}{\mathbb{P}(U \leq \varphi(Y))} \\
&= \frac{\int_a^b \mathbb{P}(U \leq \varphi(x) | Y = x) f_Y(x) dx}{\int_{\mathbb{R}} \mathbb{P}(U \leq \varphi(y) | Y = y) f_Y(y) dy} \\
&= \int_a^b \frac{\varphi(x) f_Y(x)}{\int \mathbb{P}(U \leq \varphi(y)) f_Y(y) dy} dx \\
&= \int_a^b \frac{\varphi(x) f_Y(x)}{\int \varphi(y) f_Y(y) dy} dx \\
&\stackrel{(**)}{=} \int_a^b \frac{\frac{f_X(x)}{cf_Y(x)} f_Y(x)}{\frac{1}{c}} dx \\
&= \int_a^b c \frac{1}{c} f_X(x) = \int_a^b f_X(x)
\end{aligned}$$

Im Gleichungsschritt (*) wenden wir den Satz von Bayes an. Die Umformung (**) folgt, da gilt, dass

$$\begin{aligned}
\int_{\mathbb{R}} \varphi(y) f_Y(y) dy &= \int_{\text{supp}(\mathbb{P}^Y)} \varphi(y) f_Y(y) dy = \int_{\text{supp}(\mathbb{P}^Y)} \frac{f_X(y)}{c} dy \\
&= \frac{1}{c} \int_{\text{Im}(X)} f_X(y) dy = \frac{1}{c}
\end{aligned}$$

□

Somit haben wir die Behauptung gezeigt und können zur Simulation von Zufallsvariablen neben dem Inversionsverfahren alternativ auch das „Acceptance-Rejection-Verfahren“ benutzen. Für weitere Informationen siehe auch [6].

Wie schon oben erwähnt, wollen wir im Hauptteil Pfade von stochastischen Prozessen erzeugen. Um eine mathematische Grundlage zur Betrachtung dieses Problems zu schaffen, wollen wir allerdings zunächst etwas genauer auf die Frage eingehen, was überhaupt ein stochastischer Prozess ist und eine grobe Einführung in stochastische Analysis geben, insbesondere die Bedeutung des Itô-Integrals herausstellen.

2.2 Stochastische Analysis

2.2.1 Stochastische Prozesse

Wir wollen nun zunächst einige Grundlagen über stochastische Differentialgleichungen bereitstellen. Nehmen wir an, wir befinden uns in einem zeitstetigen Modell. Bisher kennen wir (gewöhnliche) Differentialgleichungen der Form $dx(t) = a(t, x(t))dt$ oder dazu äquivalent in der Integralschreibweise $x(t) - x_0 = \int_{t_0}^t a(s, x(s))ds$ mit $x(t_0) = x_0$. Um ein realistischeres mathematisches Modell zu erhalten, in welches wir eine gewisse Zufälligkeit integrieren, wollen wir einen sogenannten „noise“-Term hinzufügen. Bei diesem kennen wir nur seine Wahrscheinlichkeitsverteilung, aber nicht die exakte Ausprägung. Die Differentialgleichung bekommt dann die Form $dX_t = a(t, X_t)dt + \sigma(t, X_t)dW_t$, wobei W_t ein Wiener Prozess ist. Wir beginnen mit den folgenden beiden Definitionen:

Definition 2.3 (Stochastischer Prozess - nach [4]).

Sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ gegeben. Ein stochastischer Prozess $X = (X_t)_{t \in T}$ ist eine Funktion $X : T \times \Omega \rightarrow \mathbb{R}^n$, wobei $X_t := X(t, \cdot)$ eine Familie von Zufallsvariablen $(X_t)_{t \in T}$ und $X(\cdot, \omega) : T \rightarrow \mathbb{R}^n$ für alle $\omega \in \Omega$ eine Realisation oder ein Pfad genannt wird. T bezeichnet dabei eine Indexmenge und repräsentiert gewöhnlich die Zeit, es gilt also normalerweise $T \subseteq [0, \infty)$.

Definition 2.4 (Wiener Prozess - nach [7]).

Ein stochastischer Prozess W , welcher adaptiert an die Filtration $(\mathcal{F}_t)_{t \in T}$ ist, heißt Wiener Prozess, wenn die folgenden Bedingungen erfüllt sind:

- (i) $W_0 = 0$ \mathbb{P} -fast sicher,
- (ii) W hat unabhängige Zuwächse, d.h. $\forall t$ und $\forall 0 \leq s < t$ ist $W_t - W_s$ unabhängig von \mathcal{F}_s ,
- (iii) W besitzt stationäre, normalverteilte Zuwächse, d.h. für $0 \leq s < t$ ist die Zufallsvariable $W_t - W_s \sim W_{t-s} \sim \mathcal{N}(0, t - s)$,
- (iv) W hat zeitstetige Pfade.

2.2.2 Das Itô-Integral

Mit den obigen Definitionen betrachten wir den Prozess $dX_t = a(t, X_t)dt + \sigma(t, X_t)dW_t$ nun in der Integralschreibweise, also als

$$X_t - X_{t_0} = \int_{t_0}^t a(s, X_s)ds + \int_{t_0}^t \sigma(s, X_s)dW_s,$$

wobei wir folglich der Frage nachgehen müssen, was genau das Integral bezüglich eines Wiener Prozesses bedeutet. Um dies zu beantworten, benötigen wir noch zwei weitere Definitionen und einen Satz:

Definition 2.5 (Beschränkte Variation - nach [8]).

Eine Funktion f auf dem Intervall heißt auf dem Intervall $[t_0, t]$ von beschränkter Variation oder auch von beschränkter totaler Variation, wenn

$$\sup \left\{ \sum_{i=1}^n |f(t_i) - f(t_{i-1})| : n \in \mathbb{N}, t_0 < t_1 < \dots < t_n \leq t \text{ ist eine Partition von } [t_0, t] \right\}$$

beschränkt ist.

Definition 2.6 (Riemann-Stieltjes-Integral).

Seien $[t_0, t] \subset \mathbb{R}$, $f, g : [t_0, t] \rightarrow \mathbb{R}^d$ gegeben mit $d \in \mathbb{N}$, für $n \in \mathbb{N}$ sei $t_0 < t_1 < \dots < t_n = t$ eine Partition von $[t_0, t]$, so dass $t_{i+1} - t_i := \Delta_i \leq \Delta \forall i \in \{0, 1, \dots, n-1\}$ und sei $\tau_i \in [t_{i-1}, t_i] \forall i \in \{1, \dots, n\}$. Dann heißt der Ausdruck

$$I := \lim_{\Delta \rightarrow 0} \sum_{i=1}^n f(\tau_i)[g(t_i) - g(t_{i-1})],$$

falls der Grenzwert existiert, das Riemann-Stieltjes-Integral von f nach g über $[t_0, t]$ und wir setzen dann

$$I = \int_{t_0}^t f(s) dg(s).$$

Der folgende Satz enthält eine Aussage über die Existenz des Riemann-Stieltjes-Integrals:

Satz 2.7 (nach [8]). Seien $f, g : [t_0, t] \rightarrow \mathbb{R}^n$ mit $f \in C[t_0, t]$ und g ist von beschränkter Variation, so existiert das Riemann-Stieltjes-Integral von f nach g über $[t_0, t]$. (Bew. siehe [8] (Satz 26))

Das Problem bei der Behandlung des Integrals bezüglich eines Wiener Prozesses ist nun, dass sich dieses weder als Riemann- noch als Riemann-Stieltjes-Integral betrachten lässt, was vielleicht zunächst vermutet werden könnte. Schreibt man nämlich für $\int_{t_0}^t \sigma(s, X_s) dW_s$ alternativ $\int_{t_0}^t \sigma(s, X_s) \frac{dW_s}{ds} ds$, ergibt sich das Problem, dass ein Wiener Prozess \mathbb{P} -fast sicher an keiner Stelle differenzierbar ist und somit der obige Ausdruck nicht existiert. Die naheliegende Interpretation des Integrals als Riemann-Stieltjes-Integral scheitert daran, dass ein Wiener Prozess Pfade von unendlicher Variation besitzt. Um dies zu zeigen, definieren wir noch die quadratische Variation:

Definition 2.8 (quadratische Variation - nach [2]).

Sei der Einfachheit wegen $t_0 = 0$ und sei $X_t : \Omega \rightarrow \mathbb{R}$ ein stetiger stochastischer Prozess. Darüber hinaus sei $0 = t_0 < t_1 < t_2 \dots < t_n = t$ für alle $n \in \mathbb{N}$ eine Partition von $[0, t]$ und $\Delta_{t_k} = t_{k+1} - t_k$. Dann ist die quadratische Variation von X_t definiert als

$$\langle X, X \rangle_t := \lim_{\Delta_{t_k} \rightarrow 0} \sum_{0 \leq t_k < t} |X_{t_{k+1}}(\omega) - X_{t_k}(\omega)|^2 \quad (\text{Limes in Wahrscheinlichkeit}),$$

sofern dieser Grenzwert existiert.

Bemerkung.

Nach [7] existiert dieser Grenzwert für stetige quadratisch integrierbare Martingale $(X_t)_t$.

Definieren wir nun $X := \sum_{0 \leq k \leq n-1} |W_{t_{k+1}} - W_{t_k}|^2 - t$, dann gilt nach dem gerade Gezeigten, dass $\mathbb{E}[X^2] = \|X\|_2^2 \leq 2\Delta_t t$ und somit folgt mit der Markov-Ungleichung (siehe Anhang) dann für alle $n \in \mathbb{N}$ und für $\Delta_t := \frac{\epsilon_n^2}{n^2} \ll 1$:

$$\mathbb{P}\left(\left|\sum_{0 \leq k \leq n-1} (W_{t_{k+1}} - W_{t_k})^2 - t\right| \geq \sqrt{2}\epsilon_n\right) \leq \frac{2\Delta_t t}{2\epsilon_n^2} = \frac{t}{n^2}$$

Daraus folgt dann wiederum:

$$\sum_{n \geq 1} \mathbb{P}\left(\left|\sum_{0 \leq k \leq n-1} (W_{t_{k+1}} - W_{t_k})^2 - t\right| > \sqrt{2}\epsilon_n\right) = \sum_{n \geq 1} \frac{t}{n^2} < \infty$$

Mit dem Borel-Cantelli-Lemma (siehe Anhang) folgt dann letztlich:

$$\mathbb{P}\left(\sum_{t_0 \leq t_k < t} (W_{t_{k+1}} - W_{t_k})^2 - t \mid > \sqrt{2}\epsilon \text{ unendlich oft}\right) = 0$$

und daraus

$$\begin{aligned} & \left|\sum_{t_0 \leq t_k < t} (W_{t_{k+1}} - W_{t_k})^2 - t\right| \leq \sqrt{2}\epsilon \text{ } \mathbb{P}\text{-fast sicher} \\ \Rightarrow & \langle W, W \rangle_t = \sum_{t_0 \leq t_k < t} (W_{t_{k+1}} - W_{t_k})^2 = t \text{ } \mathbb{P}\text{-fast sicher.} \end{aligned}$$

□

Mit Hilfe dieses Lemmas können wir nun endlich zeigen, dass ein Wiener Prozess Pfade von unendlicher Variation besitzt:

Satz 2.10. Ein Wiener Prozess besitzt Pfade von unendlicher Variation.

Beweis. Wegen der Stetigkeit der Pfade des Wiener Prozesses und aus der Abschätzung

$$\underbrace{\sum_{0 \leq t_k < t} (W_{t_{k+1}} - W_{t_k})^2}_{=t} \leq \max_k (|W_{t_{k+1}} - W_{t_k}|) \sum_{0 \leq t_k < t} |W_{t_{k+1}} - W_{t_k}|$$

sowie der Tatsache, dass $\max_k (|W_{t_{k+1}} - W_{t_k}|)$ wegen der Stetigkeit der Pfade des Wiener Prozesses \mathbb{P} -f.s. gegen 0 strebt, folgt, dass $\sum_{0 \leq t_k < t} |W_{t_{k+1}} - W_{t_k}| = \infty$, also die Behauptung. □

Wir benötigen also ein neues Integralkonzept und wollen deshalb das sogenannte Itô-Integral in drei Schritten einführen. Dazu benötigen wir einen Wiener Prozess, welcher an die Filtration $(\mathcal{F}_t)_t \subset \mathcal{F}$ adaptiert ist, so dass die Zuwächse $W_{t_{j+1}} - W_{t_j}$ unabhängig von \mathcal{F}_{t_j} sind. Nach einer weiteren Definition einer Klasse von Prozessen, für die wir das Itô-Integral definieren werden, beginnen wir mit der Konstruktion:

Definition 2.11 (nach [2]).

Eine Funktion $X(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ gehört der Klasse $L^2[0, T]$ an, falls gilt:

- (i) X ist progressiv messbar, d.h. $\forall t \in [0, T]$ gilt $(t, \omega) \mapsto X(t, \omega)$ ist $(\mathcal{B}_{[0,t]} \otimes \mathcal{F}_t)$ -messbar.
- (ii) X ist an die Filtration $(\mathcal{F}_t)_t$ adaptiert
- (iii) $\mathbb{E}[\int_0^T X(t, \omega)^2 dt] < \infty$

Zunächst sei der Prozess $X(t, \omega) \in L^2[0, T]$ ein elementarer Prozess (d.h. für alle $\omega \in \Omega$ sei der Pfad $X(\cdot, \omega) : T \rightarrow \mathbb{R}^n$ eine Treppenfunktion) mit $X(t, \omega) = X_j(\omega)$ auf $t_j \leq t < t_{j+1}$.

Sei nun $t \in [t_j, t_{j+1})$. Da X an $(\mathcal{F}_t)_t$ adaptiert ist und $X(t, \omega) = X_j(\omega)$ konstant ist auf $[t_j, t_{j+1})$, gilt $X(t, \omega) = X_j(\omega) \in \mathcal{F}_{t_j}$.

Sei nun $W_t(\omega)$ wieder ein Wiener Prozess, dann gilt

$$\begin{aligned} \mathbb{E}[X_j(\omega)(W_{t_{j+1}}(\omega) - W_{t_j}(\omega))] &\stackrel{Turm\ Eig.}{=} \mathbb{E}[\mathbb{E}(X_j(\omega)(W_{t_{j+1}}(\omega) - W_{t_j}(\omega)) | \mathcal{F}_{t_j})] \\ &\stackrel{X_j \mathcal{F}_{t_j}^-}{=} \mathbb{E}[X_j(\omega) \cdot \mathbb{E}[W_{t_{j+1}}(\omega) - W_{t_j}(\omega) | \mathcal{F}_{t_j}]] \\ &\stackrel{Unabh.}{=} \mathbb{E}[X_j(\omega) \cdot \mathbb{E}[W_{t_{j+1}}(\omega) - W_{t_j}(\omega)]] = 0. \end{aligned}$$

Definieren wir nun

$$I^T(X) = \int_0^T X(t, \omega) dW_t(\omega) := \sum_{i=0}^{n-1} X(t_i, \omega)(W_{t_{i+1}}(\omega) - W_{t_i}(\omega))$$

so folgt weiter:

Lemma 2.12 (Itô Isometrie).

Für einen beschränkten elementaren Prozess X gilt:

$$\mathbb{E}[(\int_0^T X(t, \omega) dW_t(\omega))^2] = \mathbb{E}[\int_0^T X(t, \omega)^2 dt]$$

bzw. in der Normschreibweise:

$$\| \int_0^T X(t, \omega) dW_t(\omega) \|_{\mathcal{L}_2(\mathbb{P})}^2 = \| X(t, \omega) \|_{\mathcal{L}_2(\mathbb{P} \otimes dt)}^2$$

Dabei ist $\|\cdot\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)}$ die \mathcal{L}_2 -Halbnorm bezüglich des Produktmaßes $\mathbb{P} \otimes dt$ auf $\Omega \times [0, T]$. Die intuitive Schreibweise „dt“ nutzen wir hier für das Riemann-Integral (bzw. Lebesgue-Maß). Für nähere Ausführungen siehe [3], S.46.

Beweis.

$$\begin{aligned}
\mathbb{E}[(\int_0^T X(t, \omega) dW_t(\omega))^2] &= \mathbb{E}[I^T(X)^2] \stackrel{Lin.}{=} \sum_j \mathbb{E}[X_j^2 (W_{t_{j+1}} - W_{t_j})^2] \\
&\stackrel{Turm.}{=} \sum_j \mathbb{E}[\mathbb{E}[X_j^2 (W_{t_{j+1}} - W_{t_j})^2 | \mathcal{F}_{t_j}]] \stackrel{(*)}{=} \sum_j \mathbb{E}[X_j^2 \cdot \mathbb{E}[(W_{t_{j+1}} - W_{t_j})^2]] \\
&\stackrel{(**)}{=} \sum_j \mathbb{E}[X_j^2 \cdot (Var(W_{t_{j+1}} - W_{t_j}) + \mathbb{E}[W_{t_{j+1}} - W_{t_j}]^2)] = \sum_j \mathbb{E}[X_j^2 (t_{j+1} - t_j)] \\
&= \mathbb{E}[\int_0^T X(t, \omega)^2 dt]
\end{aligned}$$

(*) gilt wegen der Unabhängigkeit der Zuwächse $W_{t_{j+1}} - W_{t_j}$ von \mathcal{F}_{t_j} und der \mathcal{F}_{t_j} -Messbarkeit von X_j . (**) folgt aus der Tatsache, dass für die Varianz einer Zufallsvariablen X gilt: $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Damit ist das Lemma bewiesen. \square

Beginnen wir nun mit der

Konstruktion des Itô-Integrals.

1. Konstruktionsschritt:

Man kann zeigen, dass für alle $X \in L^2[0, T]$ eine Folge von elementaren Prozessen $(X_n)_n \in L^2[0, T]$ existiert mit $\mathbb{E}[\int_0^T (X - X_n)^2 dt] \rightarrow 0$ für $n \rightarrow \infty$.

Für den Beweis der Aussage verweisen wir den Leser auf [2] S.27f.

Wir bemerken, dass $(X_n)_n$ damit nach [3] S.50 eine Cauchy-Folge bzgl. der \mathcal{L}_2 -Halbnorm ist.

2. Konstruktionsschritt:

Wir definieren $I_n^T(\omega) := \int_0^T X_n(t, \omega) dW(t, \omega)$ und zeigen, dass dieses gegen eine stochastische Zufallsvariable $I^T(\omega)$ in $L^2[0, T]$ konvergiert. Dazu zeigen wir, dass $I_n^T(\omega)$ eine Cauchy-Folge bzgl. der \mathcal{L}_2 -Halbnorm ist und erinnern uns an die zuvor gezeigte Itô-Isometrie.

Beweis.

$$\begin{aligned}
\mathbb{E}[(I_n^T(\omega) - I_m^T(\omega))^2] &= \mathbb{E}[(\int_0^T X_n(t, \omega) dW_t(\omega) - \int_0^T X_m(t, \omega) dW_t(\omega))^2] \\
&= \mathbb{E}[(\int_0^T X_n(t, \omega) - X_m(t, \omega) dW_t(\omega))^2] \stackrel{It\hat{o}}{=} \mathbb{E}[\int_0^T (X_n(t, \omega) - X_m(t, \omega))^2 dt] \xrightarrow{1} 0
\end{aligned}$$

\square

3. Konstruktionsschritt:

Wir setzen zuletzt $I^T(\omega) := \lim_{n \rightarrow \infty} I_n(\omega)$ (mit Grenzwert in L^2) und nennen dies das Itô-Integral.

Abschließend sei noch angemerkt, dass, wenn man bei der Definition des Integrals bzgl. eines Wiener Prozesses für elementare Prozesse anstatt die Funktion an der

linken Intervallgrenze des Intervalls $[t_i, t_{i+1}]$ auszuwerten - also $X(t_i, \omega)$ - diese in der Mitte des Intervalls ausgewertet - also $X(\frac{t_{i+1}+t_i}{2}, \omega)$ -, man statt des Itô-Integrals das sogenannte Stratonovich-Integral erhält. Offensichtlich macht es also bei der Definition von stochastischen Integralen im Gegensatz zum Riemann-Stieltjes-Integral einen Unterschied, an welcher Stelle man die Funktion auswertet. Der Grund, warum sich das Itô-Integral in der Finanzmathematik „durchgesetzt“ hat, liegt zum einen daran, dass der Kurs des underlying assets zu Beginn des Zeitraums $[t_i, t_{i+1}]$ nur in t_i , also am linken Intervallrand, bekannt ist, zum anderen an folgendem Satz:

Satz 2.13 (Martingaleigenschaft des Itô-integrals). Für jede Funktion $\sigma(t, \omega)$ mit $\sigma \in L^2[0, T] \forall T$ ist der Prozess $X = (X_t)_t := \int_0^t \sigma(s) dW_s$ als Funktion von t ein Martingal.

Beweis. [Beweisskizze] Sei X weiterhin an die Filtration $(\mathcal{F}_t)_t$ adaptiert und $s < t$. Dann gilt:

$$\begin{aligned} \mathbb{E}[X_t | \mathcal{F}_s] &= \mathbb{E}\left[\int_0^t \sigma(u) dW_u \mid \mathcal{F}_s\right] = \mathbb{E}\left[\int_0^s \sigma(u) dW_u \mid \mathcal{F}_s\right] + \mathbb{E}\left[\int_s^t \sigma(u) dW_u \mid \mathcal{F}_s\right] \\ &= \int_0^s \sigma(u) dW_u + \mathbb{E}\left[\int_s^t \sigma(u) dW_u\right] = \int_0^s \sigma(u) dW_u = X_s \end{aligned}$$

Die dritte Gleichheit ist erfüllt, weil das Integral $\int_0^s \sigma(u) dW_u$ \mathcal{F}_s -messbar ist. Die vorletzte Gleichung gilt, da $\mathbb{E}[\int_s^t \sigma(u) dW_u | \mathcal{F}_s] = 0$ ist. Diese Tatsachen kann man beweisen, indem man sie zunächst für elementare Prozesse σ beweist und dann wie bei der Konstruktion des Itô-Integrals über den Grenzwert in L^2 auf allgemeine L^2 -Funktionen übergeht. Für einen Beweis des Satzes siehe beispielsweise [2] oder [7]. \square

Satz 2.14. Die Itô-Isometrie gilt auch für allgemeine Prozesse in $L^2[0, T]$.

Beweis. Sei X_n wie bei der Konstruktion des Itô-Integrals eine Folge elementarer Prozesse mit $\|X_n\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)} \rightarrow \|X\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)}$ und $\|I(X_n) - I(X)\|_{\mathcal{L}_2(\mathbb{P})} \rightarrow 0$. Dann folgt, dass $\|I(X_n)\|_{\mathcal{L}_2(\mathbb{P})} \rightarrow \|I(X)\|_{\mathcal{L}_2(\mathbb{P})}$. Andererseits gilt wegen der Itô-Isometrie für elementare Prozesse $\|I(X_n)\|_{\mathcal{L}_2(\mathbb{P})} = \|X_n\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)}$, also insgesamt:

$$\|I(X)\|_{\mathcal{L}_2(\mathbb{P})} \leftarrow \|I(X_n)\|_{\mathcal{L}_2(\mathbb{P})} = \|X_n\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)} \rightarrow \|X\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)}.$$

Wegen der Tatsache, dass L_2 ein Hilbert-Raum ist (siehe u.a.[3]), folgt, dass die Grenzwerte übereinstimmen, so dass gilt:

$$\mathbb{E}\left[\int_0^T X_t^2 dt\right] = \|X\|_{\mathcal{L}_2(\mathbb{P} \otimes dt)}^2 = \|I(X)\|_{\mathcal{L}_2(\mathbb{P})}^2 = \mathbb{E}\left[\left(\int_0^T X_t dW_t\right)^2\right],$$

also die Itô-Isometrie. \square

Zum Abschluss der Diskussion des Itô-integrals weisen wir noch darauf hin, dass sich das Integral nach [2] auch auf eine Klasse mit schwächeren Bedingungen als denen aus Definition 2.11 erweitern lässt:

Bemerkung 2.15. Sei $\mathcal{W}^2[0, T]$ die Klasse aller Prozesse $X(t, \omega) \in \mathbb{R}$, die folgende Bedingungen erfüllen:

- (i) X ist progressiv messbar, d.h. $\forall t \in [0, T]$ gilt $(t, \omega) \mapsto X(t, \omega)$ ist $(\mathcal{B}_{[0,t]} \otimes \mathcal{F}_t)$ -messbar,
- (ii)' es existiert eine monoton wachsende Familie von σ -Algebren $(\mathcal{H}_t)_{t \geq 0}$, so dass
 - (a) W_t ein Martingal bzgl. \mathcal{H}_t ist und
 - (b) X_t \mathcal{H}_t -adaptiert ist,
- (iii)' $\mathbb{P}[\int_0^T X(t, \omega)^2 dt < \infty] = 1$.

Dann können wir das Itô-Integral analog zu oben mit der Ausnahme, dass das Integral diesmal als Grenzwert in Wahrscheinlichkeit anstatt als Grenzwert in L^2 existiert, auch für $X \in \mathcal{W}^2[0, T]$ definieren.

2.2.3 Die Itô-Formel

Da wir nun geklärt haben, was ein Integral bezüglich eines Wiener Prozesses und damit was ein stochastischer Prozess ist, wollen wir nun noch die sehr bekannte Itô-Formel angeben, welche man als Kettenregel der stochastischen Analysis ansehen kann:

Satz 2.16 (Itô-Formel).

Sei X_t ein stochastischer Prozess der Form $X_t = X_0 + \int_0^t \mu(s, \omega) ds + \int_0^t \sigma(s, \omega) dW_s$, wobei μ und σ adaptierte Prozesse seien mit

$$\mathbb{P}[\int_0^t |\mu(s, \omega)| ds < \infty \forall t \geq 0] = 1 \text{ und } \mathbb{P}[\int_0^t \sigma(s, \omega)^2 ds < \infty \forall t \geq 0] = 1.$$

Nachfolgend schreiben wir der Einfachheit halber nur μ und σ für $\mu(s, \omega)$ bzw. $\sigma(s, \omega)$.

Ist dann f eine $\mathcal{C}^{1,2}$ Funktion auf $[0, \infty) \times \mathbb{R}$ und der Prozess $Z(t) := f(t, X_t)$, so besitzt Z das stochastische Differential gegeben durch:

$$\begin{aligned} dZ(t) &= df(t, X_t) = \frac{\partial f(t, X_t)}{\partial t} dt + \frac{\partial f(t, X_t)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f(t, X_t)}{\partial x^2} d\langle X, X \rangle_t \\ &= \left(\frac{\partial f(t, X_t)}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma \frac{\partial f}{\partial x} dW(t) \end{aligned}$$

Für die zweite Gleichheit machen wir folgende

Bemerkung 2.17. Es gilt $d\langle X, X \rangle_t = \sigma^2 dt$

Beweis.

Sei $0 = t_0 < t_1 < \dots < t_n = t$ für ein $n \in \mathbb{N}$ eine Partition von $[0, t]$ mit $\max_i (t_{i+1} - t_i) := \Delta_t$. Da X ein L^2 -Martingal ist, existiert dessen quadratische

Variation und es gilt:

$$\begin{aligned}
\langle X, X \rangle_t &= \langle X_0 + \int_0^\cdot \mu ds + \int_0^\cdot \sigma dW_s, X_0 + \int_0^\cdot \mu ds + \int_0^\cdot \sigma dW_s \rangle_t \\
&= \lim_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} | X_0 + \int_0^{t_{i+1}} \mu ds + \int_0^{t_{i+1}} \sigma dW_s - (X_0 + \int_0^{t_i} \mu ds + \int_0^{t_i} \sigma dW_s) |^2 \\
&= \lim_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} \left(\int_0^{t_{i+1}} \mu ds - \int_0^{t_i} \mu ds \right)^2 + 2 \left[\left(\int_0^{t_{i+1}} \mu ds - \int_0^{t_i} \mu ds \right) \right. \\
&\quad \left. \left(\int_0^{t_{i+1}} \sigma dW_s - \int_0^{t_i} \sigma dW_s \right) \right] + \left(\int_0^{t_{i+1}} \sigma dW_s - \int_0^{t_i} \sigma dW_s \right)^2 \\
&\leq \langle \int_0^\cdot \mu ds, \int_0^\cdot \mu ds \rangle_t + 2 \lim_{\Delta t \rightarrow 0} \underbrace{\left(\sup_j \left| \int_0^{t_{j+1}} \sigma dW_s - \int_0^{t_j} \sigma dW_s \right| \right)}_{\rightarrow 0, \text{ da It\^o-Integral stetig in } t} \cdot \underbrace{\sum_{i=0}^n \int_{t_i}^{t_{i+1}} \mu ds}_{< \infty \text{ f.s.}} \\
&\quad + \langle \int_0^\cdot \sigma dW_s, \int_0^\cdot \sigma dW_s \rangle_t = \langle \int_0^\cdot \sigma dW_s, \int_0^\cdot \sigma dW_s \rangle_t
\end{aligned}$$

Im letzten Beweisschritt nutzen wir die Tatsache, dass ein Riemann-Integral differenzierbar ist und somit quadratische Variation „=0“ hat. Ersetzen wir im vorletzten Schritt das Supremum durch das Infimum, so erhalten wir analog

$$\langle X, X \rangle_t \geq \langle \int_0^\cdot \sigma dW_s, \int_0^\cdot \sigma dW_s \rangle_t.$$

Insgesamt folgt also die Gleichheit. Deshalb erhalten wir

$$d\langle X, X \rangle_t = d\langle \int_0^\cdot \sigma dW_s, \int_0^\cdot \sigma dW_s \rangle_t.$$

Mit der Linearität des Itô-Integrals (siehe [2]) und der Itô-Isometrie erhalten wir letztlich:

$$\begin{aligned}
d\langle \int_0^\cdot \sigma dW_s, \int_0^\cdot \sigma dW_s \rangle_t &\stackrel{Lin.}{=} \stackrel{Isom.}{=} d \lim_{\Delta t \rightarrow 0} \sum_i \left| \int_{t_i}^{t_{i+1}} \sigma^2 ds \right| \\
&= d \int_0^t \sigma^2 ds = \sigma^2 dt,
\end{aligned}$$

wobei wir für die letzte Gleichheit den Hauptsatz der Diff.-und Int.-rechnung verwenden. \square

Für einen vollständigen Beweis der Itô-Formel siehe [7]. Wir finden dort den Bew. der Formel für stetige Semimartingale, also gilt diese insb. für Itô-Prozesse X .

2.2.4 Spezielle SDEs und „strong solutions“

Zum Abschluss der Einführung stochastischer Differentialgleichungen wollen wir noch zwei Spezialfälle betrachten sowie kurz die Existenz von Lösungen von SDEs diskutieren. Beginnen wir mit dem sogenannten

Ornstein-Uhlenbeck Prozess.

Die zugehörige SDE lautet:

$$dX_t = -aX_t dt + b dW_t, \text{ wobei } a, b > 0 \text{ hier konstant seien.}$$

Dann existiert eine explizite Lösung, und zwar: $X_t = e^{-at}(X_0 + \int_0^t e^{as} b dW_s)$.

Als weiteres und wichtiges Beispiel betrachten wir noch die

geometrische Brownsche Bewegung.

Die zugehörige SDE lautet:

$$dX_t = aX_t dt + bX_t dW_t, \text{ wobei hier } a, b \in \mathbb{R} \text{ ebenfalls konstant seien.}$$

Auch in diesem Fall existiert eine explizite Lösung, und zwar: $X_t = X_0 e^{(a - \frac{1}{2}b^2)t + bW_t}$.

Mithilfe der Itô-Isometrie können wir die Lösung leicht herleiten:

Definiere dazu $f(t, X_t) := \ln(X_t)$. Dann folgt:

$$\begin{aligned} df(t, X_t) &= \underbrace{\left(\frac{\partial f(t, X_t)}{\partial t}\right)}_{=0} + aX_t \underbrace{\left(\frac{\partial f(t, X_t)}{\partial x}\right)}_{=X_t^{-1}} + \frac{1}{2}(bX_t)^2 \underbrace{\left(\frac{\partial^2 f(t, X_t)}{\partial x^2}\right)}_{=-X_t^{-2}} dt + bX_t \underbrace{\left(\frac{\partial f(t, X_t)}{\partial x}\right)}_{=X_t^{-1}} dW_t \\ &= \left(a - \frac{1}{2}b^2\right)dt + b dW_t \end{aligned}$$

$$\begin{aligned} \stackrel{\text{Integral-}}{\text{Schreibweise}} \ln(X_t) &= \ln(X_0) + \int_0^t \left(a - \frac{1}{2}b^2\right)ds + \int_0^t b dW_s \\ &= \ln(X_0) + \left(a - \frac{1}{2}b^2\right)t + bW_t \\ \Rightarrow X_t &= X_0 e^{(a - \frac{1}{2}b^2)t + bW_t} \end{aligned}$$

Zuletzt geben wir noch den Erwartungswert von X_t an. Nach [2] gilt:

$$\mathbb{E}[X_t] = \mathbb{E}[X_0]e^{rt}. \tag{2.1}$$

Betrachten wir zum Abschluss des Kapitels noch den allgemeinen und häufigen Fall, dass keine explizite Lösung der stochastischen Differentialgleichung

$$dX_t = \underbrace{\mu(t, X_t)}_{\text{Drift}} dt + \underbrace{\sigma(t, X_t)}_{\text{Diffusion}} dW_t, \quad t \in [0, T]$$

existiert. Die folgende Definition und der folgende Satz (beide nach [6]) geben Aufschluss über die Existenz einer Lösung.

Definition 2.18 (Strong Solution - nach [6]).

Eine sogenannte „strong solution“ der obigen stochastischen Differentialgleichung auf dem Intervall $[0, T]$ ist ein stochastischer Prozess X mit

$$X_t = X_0 + \int_0^t \mu(u, X_u) du + \int_0^t \sigma(u, X_u) dW_u, \quad t \in [0, T],$$

welcher die Bedingungen $\mathbb{P}(\int_0^T |\mu(t, X_t)| dt < \infty) = 1$ und $\mathbb{P}(\int_0^T \|\sigma(t, X_t)\|^2 dt < \infty) = 1$ erfüllt.

Der folgende und abschließende Satz gibt uns eine Aussage über die Existenz und Eindeutigkeit einer „strong solution“ einer stochastischen Differentialgleichung:

Satz 2.19 (Existenz und Eindeutigkeit).

Unter der Annahme, dass $\mathbb{E}(\|X_0^2\|) < \infty$ gilt und der Existenz einer Konstante K mit

$$\|\mu(t, x) - \mu(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\| \quad (\text{Lipschitz-Bedingung})$$

$$\text{und } \|\mu(t, x)\| + \|\sigma(t, x)\| \leq K(1 + \|x\|) \quad (\text{Lineare Wachstums-Bedingung})$$

$\forall t \in [0, T]$ und $\forall x, y \in \mathbb{R}^d$ existiert eine „strong solution“ X . Sie ist eindeutig in dem Sinne, dass gilt $\mathbb{P}(X_t = \tilde{X}_t \forall t \in [0, T]) = 1$ und alle \tilde{X} mit \tilde{X} ist Lösung der SDE) = 1. Des weiteren erfüllt die Lösung dann für alle $t \in [0, T]$ die Abschätzung

$$\mathbb{E}[\|X_t\|^2] < \infty.$$

3 Dynamische Monte Carlo Methoden

Wie wir schon in der Einleitung erwähnt haben, ist es das Ziel von Monte Carlo Methoden einen Erwartungswert über eine Zufallsvariable bzw. einen stochastischen Prozess zu schätzen. Mathematisch bedeutet dies, dass man Erwartungswerte der Form

$$\mathbb{E}[e^{-rT}V(T)|\mathcal{F}_t],$$

wobei V einen stochastischen Prozess, also $V(T)$ eine Zufallsgröße, r die risikolose Zinsrate und \mathcal{F}_t für alle t eine σ -Algebra darstellen. Gibt $V(T)$ den Wert einer Option zum Zeitpunkt T an, so kann man mit dem obigen Erwartungswert den Preis der Option mit Laufzeit T zum Zeitpunkt t berechnen.

Betrachten wir eine Europäische Plain-Vanilla Call-Option, so beläuft sich dieser Wert auf

$$C_t = \mathbb{E}[e^{-r(T-t)}\max\{S_T - K, 0\}|\mathcal{F}_t],$$

wobei S_T den Wert des underlying assets zum Zeitpunkt T und K den Strike-Preis der Option angeben. Die Filtration $(\mathcal{F}_t)_t$ ist die vom Wiener Prozess (als zufällige Komponente des stochastischen Prozesses, welcher das underlying asset beschreibt) erzeugte Filtration. In diesem Abschnitt werden wir sehen, wie wir das underlying asset mathematisch modellieren und dessen Pfad simulieren können. Dazu benötigen wir dynamische Monte Carlo Methoden. Wir gehen bei der Modellierung davon aus, dass das underlying asset durch eine stochastische Differentialgleichung beschrieben wird.

3.1 Diskretisierungsverfahren

Bei vielen für den Finanzmarkt wichtigen stochastischen Differentialgleichungen (SDEs) ist keine explizite Lösung verfügbar. In diesem Fall können wir den Prozess, welcher durch die SDE beschrieben wird, diskretisieren, um dann eine exakte Lösung des diskreten Prozesses, welcher die SDE

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) \tag{3.1}$$

approximiert, zu erhalten. Diese Approximation erhält man, indem man die SDE geeignet diskretisiert. Dazu teilt man das zu betrachtende Intervall in N ($\in \mathbb{N}$) Intervalle und simuliert den Prozess an den so entstehenden $N+1$ Stellen. Die bekanntesten Methoden für diese Diskretisierung sind das Euler-(Maruyama-)Verfahren sowie das Milstein-Verfahren. Beide werden nun kurz vorgestellt und es wird auf Konvergenzarten als Vergleichskriterium eingegangen.

Definition 3.1 (Euler-Maruyama-Verfahren).

Sei $(X_t)_{t \geq 0}$ ein stochastischer Prozess, welcher durch die obigen Dynamics gegeben ist. Dann definiert die Approximation $(Y_n)_n$ gegeben durch

$$Y_{n+1} = Y_n + \mu(t_n, Y_n) \cdot \Delta_n + \sigma(t_n, Y_n) \cdot \Delta W_n \quad \forall 1 \leq n \leq N$$

mit $\Delta_n = t_{n+1} - t_n$ und $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ (Zuwachs des Wiener Prozesses W) die Diskretisierung von X mittels des Euler-Maruyama-Verfahrens.

Betrachtet man die Konvergenzordnung des Euler-Verfahrens aus Sicht der Taylorentwicklung, so kann man nach [6] sehen, dass diese den Diffusionsterm nur mit Ordnung $\mathcal{O}(\sqrt{\Delta})$ entwickelt, während sie es beim Driftterm mit Ordnung $\mathcal{O}(\Delta)$ kann. Deshalb ist es die Idee des Milstein-Verfahrens einen Term höherer Ordnung aus der Taylorentwicklung zum Euler-Verfahren hinzuzufügen, nämlich den Term $\frac{1}{2}\sigma\sigma' \cdot ((\Delta W_n)^2 - \Delta_n)$. Wir erhalten dann das folgende

Definition 3.2 (Milstein-Verfahren).

$$Y_{n+1} = Y_n + \mu \cdot \Delta_n + \sigma \cdot \Delta W_n + \frac{1}{2}\sigma\sigma' \cdot ((\Delta W_n)^2 - \Delta_n) \quad \forall 1 \leq n \leq N$$

mit $\Delta_n = t_{n+1} - t_n$ und $\Delta W_n = W_{t_{n+1}} - W_{t_n}$.

Um zu zeigen, ob und in welchem Sinne es sich tatsächlich um eine Verbesserung des Euler-Verfahrens handelt, betrachten wir je ein Kriterium für die pfadweise Nähe des diskretisierten Prozesses zum zeitstetigen Prozess, sowie für die Nähe der jeweiligen Verteilungen, nämlich das starke und das schwache Konvergenzverhalten, welche wie in [4] angeben wollen:

Definition 3.3 (schwache Konvergenz).

Eine allgemeine diskrete Approximation $(Y_n)_n$ mit maximaler Schrittweite Δ ($:= \max\{\Delta_n : 1 \leq n \leq N\}$) konvergiert schwach gegen den stochastischen Prozess $(X_t)_{t \geq 0}$ (Lösung der obigen SDE 3.1) zur Zeit T , für $\Delta \downarrow 0$, in Bezug auf eine Klasse \mathcal{C} von Testfunktionen $g : \mathbb{R}^d \rightarrow \mathbb{R}$, falls gilt:

$$\lim_{\Delta \downarrow 0} |\mathbb{E}(g(X_T)) - \mathbb{E}(g(Y(T)))| = 0 \quad \forall g \in \mathcal{C}$$

Definition 3.4 (starke Konvergenz).

Eine allgemeine diskrete Approximation $(Y_n)_n$ mit maximaler Schrittweite Δ ($:= \max\{\Delta_n : 1 \leq n \leq N\}$) konvergiert stark gegen den stochastischen Prozess $(X_t)_{t \geq 0}$ (Lösung der obigen SDE 3.1) zur Zeit T , für $\Delta \downarrow 0$, falls gilt:

$$\lim_{\Delta \downarrow 0} \mathbb{E}|X_T - Y(T)| = 0.$$

Eine alternative Bedingung für starke Konvergenz, die man häufig findet (siehe bspw. [6]), lautet:

$$\lim_{\Delta \downarrow 0} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - Y(t)| \right] = 0.$$

Um nun Diskretisierungsverfahren zu vergleichen, betrachtet man die jeweiligen Konvergenzraten:

Definition 3.5 (Konvergenzrate der starken Konvergenz).

Eine diskrete Approximation $(Y_n)_n$ konvergiert stark mit Rate $\gamma \geq 0$ zur Zeit T , falls eine positive Konstante C , welche nicht von Δ abhängt, und ein δ_0 existieren, so dass

$$\epsilon(\Delta) = \mathbb{E}(|X_T - Y(T)|) \leq C\Delta^\gamma \quad \forall \Delta \in (0, \delta_0).$$

Analog dazu definiert man ebenfalls die Konvergenzrate der schwachen Konvergenz.

Man kann nun feststellen, was wir hier nur angeben möchten, dass das Euler-Verfahren schwach mit Rate 1 und stark mit Rate $\frac{1}{2}$ konvergiert, wohingegen das Milstein-Verfahren sowohl schwach als auch stark mit Rate 1 konvergiert. Es handelt sich beim Milstein-Verfahren folglich im Vergleich zum Euler-Verfahren tatsächlich um eine Verbesserung der (starken) Konvergenzrate.

Bei der Berechnung von Preisen von Derivaten im Finanzmarkt sind offensichtlich schwache Fehlerkriterien relevanter, da Preise oft als Erwartungswerte von Funktionen angewendet auf stochastische Prozesse berechnet werden. Deshalb möchte man sicherstellen, dass von X_t berechnete Preise denen von Y_n möglichst genau entsprechen.

Betrachtet man Pfade von stochastischen Prozessen, so ist ebenfalls das starke Konvergenzverhalten von großer Bedeutung. Dieses benötigt man folglich besonders bei den dynamischen Monte Carlo Methoden, bei welchen wir (die kompletten) Pfade betrachten müssen, um beispielsweise Preise von Asiatischen Optionen zu bestimmen. Für den Rest des Kapitels wollen wir uns grob an der Darstellung in [1] orientieren.

3.2 Simulation einer Brownschen Bewegung

Beginnen wir nun mit der ersten Simulation eines Pfades eines stochastischen Prozesses, nämlich der Brownschen Bewegung

$$dX_t = dW_t.$$

Haben wir den zeitstetigen Prozess $(X_t)_t$ mithilfe einer Diskretisierungsmethode durch den Vektor $X^{(0)}, \dots, X^{(N)}$ approximiert, welcher die Werte $X_{[0,T]}$ an festen Stellen $0 = t_0 < t_1 < \dots < t_N = T$ repräsentiert, so können wir uns bei der Simulation dann auf die so entstehenden Zufallsvariablen beschränken. Falls der komplette Pfad benötigt wird, können wir diesen durch Interpolation erhalten. Eine naheliegende (und vor allem einfache) Möglichkeit besteht in der linearen Verbindung der simulierten Punkte.

Wir sehen folglich, dass wir, um zeitstetige Prozesse zu simulieren, ebenfalls Zufallsvariablen simulieren können müssen. Neben der Möglichkeit Samples von Zufallsvariablen über die Erzeugung eines Samples der Gleichverteilung (Inversionsverfahren) oder durch das „Acceptance-Rejection-Verfahren“ zu erhalten, benötigen wir bei der Simulation eines Wiener Prozesses zufällig erzeugte Samples der Standardnormalverteilung. Diese können wir beispielsweise mit der Box-Müller Methode (siehe dazu u.a. [6]) erzeugen. Die Simulation fällt auch unter die Kategorie der statischen Monte Carlo Methoden. Deshalb werden wir darauf nicht weiter eingehen und annehmen, dass standardnormalverteilte Samples schnell und einfach erzeugt werden können - in Matlab beispielsweise mit dem Befehl „randn“. Für die Simulation von stochastischen Prozessen benötigen wir die Fähigkeit einen Wiener Prozess bzw. eine Brownsche Bewegung simulieren zu können, denn diese/r bildet die Komponente in stochastischen Differentialgleichungen, in der die

Zufälligkeit steckt. Wir approximieren den Wiener Prozess $W = (W_t)_t$ beispielsweise mittels des Euler-Verfahrens. Statt die Punkte linear zu verbinden, betrachten wir einen auf endlich vielen aufeinanderfolgenden Intervallen konstanten Prozess. Wählen wir bei der Diskretisierung eine konstante Schrittweite Δ_t , so erhalten wir den Vektor der Diskretisierungspunkte als $(W_0, W_{\Delta_t}, \dots, W_{N\Delta_t})$. Jeder Brownsche Zuwachs $W_{i\Delta_t} - W_{(i-1)\Delta_t}, 1 \leq i \leq N$, ist nach Definition $\mathcal{N}(0, \Delta_t)$ verteilt. Das heißt der Vektor $(W_0, W_{\Delta_t}, \dots, W_{N\Delta_t})$ ist dann genauso verteilt wie $(0, \sqrt{\Delta_t}Z_1, \dots, \sqrt{\Delta_t}(Z_1 + \dots + Z_N))$, wobei $Z_i \sim \mathcal{N}(0, 1)$ für alle $0 \leq i \leq N$ eine standardnormalverteilte Zufallsgröße sei. Dies ist durch die Tatsache zu erklären, dass Summen von normalverteilten Zufallsgrößen wieder normalverteilte Zufallsgrößen sind. Es gilt also für alle $0 \leq i \leq N$

$$Z_1 + \dots + Z_{i\Delta_t} \sim \mathcal{N}(0, \underbrace{1 + \dots + 1}_{i\text{-mal}}) = \mathcal{N}(0, i) \text{ und daher } \sqrt{\Delta_t}(Z_1 + \dots + Z_i) \sim \mathcal{N}(0, \Delta_t i).$$

Deshalb können wir W_t approximieren durch $W'_t := W_{\lfloor t/\Delta_t \rfloor \Delta_t} \sim \sqrt{\Delta_t}(Z_1 + \dots + Z_{\lfloor t/\Delta_t \rfloor})$, wobei $\lfloor x \rfloor$ dem ganzzahligen Anteil von x entspricht und wir in diesem Fall annehmen, dass W'_t auf dem Intervall $[(i-1)\Delta_t, i\Delta_t)$ für alle $1 \leq i \leq N$ konstant ist.

Im folgenden Abschnitt wollen wir die Simulation von Pfaden eines Wiener Prozesses nutzen, um Pfade der geometrischen Brownschen Bewegung, welche das underlying asset im Black-Scholes Modell beschreiben, zu simulieren.

3.3 Simulation einer geometrischen Brownschen Bewegung

In einigen Spezialfällen ist es möglich, die Lösung einer stochastischen Differentialgleichung explizit anzugeben, d.h. es existiert ein analytisches Funktional G mit $X_t = G(t, W_{[0,t]})$. In diesem Fall können wir statt der SDE auch die gegebene Lösung diskretisieren. Der Algorithmus ergibt sich dann wie folgt:

Algorithmus 3.6 (Exact Solution). 1. Lege eine natürliche Zahl N fest, welche die Anzahl der Diskretisierungsstellen festlegt.

2. Erzeuge für $i = 1, \dots, N$ Samples der Brownschen Bewegung $W(t_i)$ und setze $X_i = G(t_i, \{W_{t_1}, \dots, W_{t_i}\})$

3. Output: $(X_{\lfloor t/\Delta_t \rfloor}, 0 \leq t \leq T)$ als Sample-Pfad des Prozesses X auf $[0, T]$.

Als Beispiel schauen wir uns die geometrische Brownsche Bewegung $dS_t = rS_t dt + \sigma S_t dW_t$, deren Lösung durch

$$S_t = G(t, W_t) := S_0 \cdot \exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma W_t\right)$$

gegeben ist, näher an.

Der große Vorteil, die geometrische Brownsche Bewegung über die exakte Lösung anstatt über die Differentialgleichung zu simulieren, besteht darin, dass der simulierte Prozess - vorausgesetzt $S_0 > 0$ - wegen des Exponentialterms nie negative Werte annehmen kann. Dies ist eine sehr wichtige Eigenschaft, denn in der Realität kann der Wert einer Aktie natürlich ebenfalls nicht kleiner als 0 sein.

Das Problem, (mithilfe einer ausreichend kleinen Schrittgröße) aufpassen zu müssen, dass der Pfad des Prozesses immer nichtnegativ bleibt, wird somit also von vornherein vermieden. Simuliert man den Prozess S_t mit

$$S_{i+1} = S_i \cdot \exp\left(\left(r - \frac{\sigma^2}{2}\right)\Delta_t + \sigma\sqrt{\Delta_t}Z_{i+1}\right)$$

und bezeichnet Z_i dabei für alle i wieder ein Sample der Standardnormalverteilung, so erhält man einen Samplepfad, mit welchem man zum Beispiel den Preis einer Option im Black-Scholes Modell berechnen kann. Möchte man die S_{i+1} konkret als $G(t_i, \{W_{t_1}, \dots, W_{t_i}\})$ schreiben, so lautet die Diskretisierung

$$S_{i+1} = S_0 \cdot \exp\left(\left(r - \frac{\sigma^2}{2}\right) \underbrace{\Delta_t \cdot (i+1)}_{t_{i+1}} + \sigma \underbrace{\sqrt{\Delta_t}(Z_1 + \dots + Z_{i+1})}_{W_{t_{i+1}}}\right).$$

Die Äquivalenz zur obigen Schreibweise lässt sich sehr leicht durch Induktion zeigen.

Als einfache Anwendung betrachten wir nun eine Europäische Plain-Vanilla Call-Option. Der Wert C_t der Option zum Zeitpunkt $t < T$ beläuft sich bekanntlich auf

$$C_t = \mathbb{E}[e^{-r(T-t)} \max\{S_T - K, 0\} | \mathcal{F}_t],$$

wobei S_T den Wert des underlying assets zum Zeitpunkt T und K den Strike-Preis der Option angeben. Die Filtration $(\mathcal{F}_t)_t$ ist wie oben die vom Wiener Prozess erzeugte Filtration. D.h. wir können \mathcal{F}_t auch durch S_t ersetzen; denn ist der Pfad des Wiener Prozesses bis zum Zeitpunkt t bekannt, so ist damit auch der Pfad des underlying assets $(S_s)_{0 \leq s \leq t}$ bekannt. Für den Wert der Option in T und den weiteren Verlauf des underlying assets ist aber lediglich der Wert S_t im Zeitpunkt t relevant (Markov-Eigenschaft). Da wir der Einfachheit halber für die weitere Arbeit voraussetzen wollen, dass wir nur den Wert der Option zum Zeitpunkt 0 berechnen wollen, wir also immer $t=0$ betrachten, beläuft sich der Preis einer Call-Option dann auf

$$C = \mathbb{E}[e^{-rT} \max\{S_T - K, 0\} | S_0],$$

wobei wir in Zukunft einfach nur

$$C = \mathbb{E}[e^{-rT} \max\{S_T - K, 0\}]$$

schreiben werden. Um das Verfahren auch numerisch anzuwenden, haben wir für unterschiedliche $n \in \mathbb{N}$ jeweils n Pfade der geometrischen Brownschen Bewegung mit dem obigen Algorithmus simuliert und den zugehörigen Call-Preis zum Zeitpunkt $t=0$

$$\mathbb{E}[e^{-rT} \cdot \max\{S_T - K, 0\}] \approx \frac{1}{n} \left(\sum_{j=1}^n e^{-rT} \cdot \max\{S_T^{(j)} - K, 0\} \right)$$

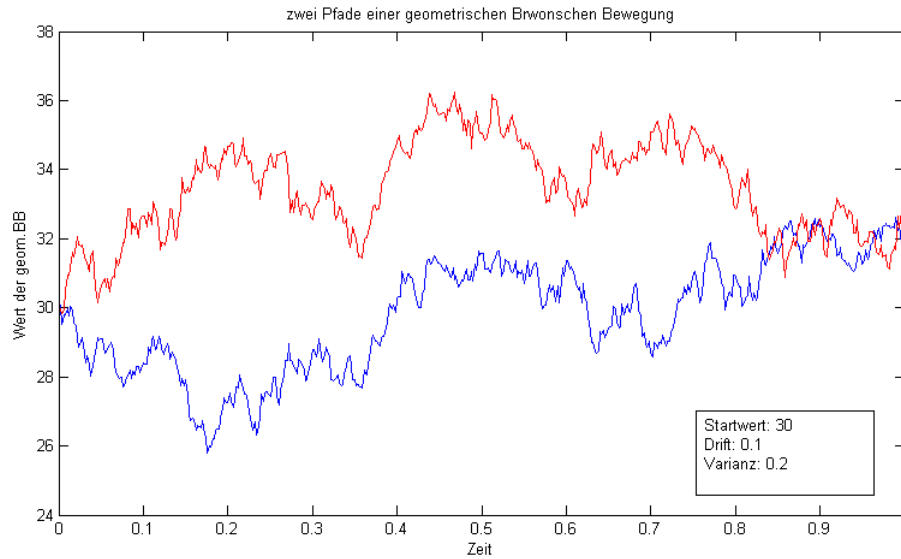


Abbildung 1: zwei approximierte Pfade einer geometrischen Brownschen Bewegung mit Drift 0.1 und Volatilität 0.2

berechnet, wobei $S_T^{(j)}$ den Wert der j -ten Simulation des Prozesses S_t zum Zeitpunkt T , K den Strikepreis und r die risikolose Zinsrate bezeichnen. Als Ergebnis erhalten wir den durchschnittlichen Call-Preis in Abhängigkeit von n . Wir können sehen, dass je mehr Pfade wir simulieren, sich das Ergebnis im Durchschnitt immer weiter dem mit der Black-Scholes Formel berechneten Call-Preis annähert. Dies ist genau die Aussage des Gesetzes der großen Zahlen, auf welchem die Monte Carlo Methoden, wie in der Einleitung erläutert, basieren. Natürlich können (und werden) die Abweichungen mit Zunahme der Anzahl der Pfade auch mal wieder ansteigen, jedoch steigt die Wahrscheinlichkeit, dass die Abweichung kleiner wird mit der Anzahl der simulierten Pfade an. In der Abbildung 2 können wir diesen Effekt gut erkennen. Selbstverständlich besitzt die Grafik für den allgemeinen Fall keine Aussagekraft, jedoch ist sie einigermaßen repräsentativ.

Betrachten wir in diesem Beispiel zuletzt die Fehlerentwicklung in Bezug auf die Anzahl der simulierten Pfade. Die Ergebnisse in Abbildung 2 spiegeln die durchschnittlichen Ergebnisse bzgl. der Fehler bzw. Abweichungen zum Black-Scholes-Preis recht gut wider. Wir können sehen, dass der relative Fehler hier ab 20000 simulierten Pfaden immer kleiner als 1% des Call-Preises ist. Eine Tabelle (Tabelle 1) der Werte und Approximationsfehler befindet sich im Anhang.

3.4 Jump Prozesse

So wichtig das Black-Scholes Modell für die Finanzmathematik auch ist, reicht uns der Wiener Prozess als zufällige Komponente des underlying assets nicht immer aus, um ein gutes Modell der Realität zu erstellen. Eine Möglichkeit das Modell

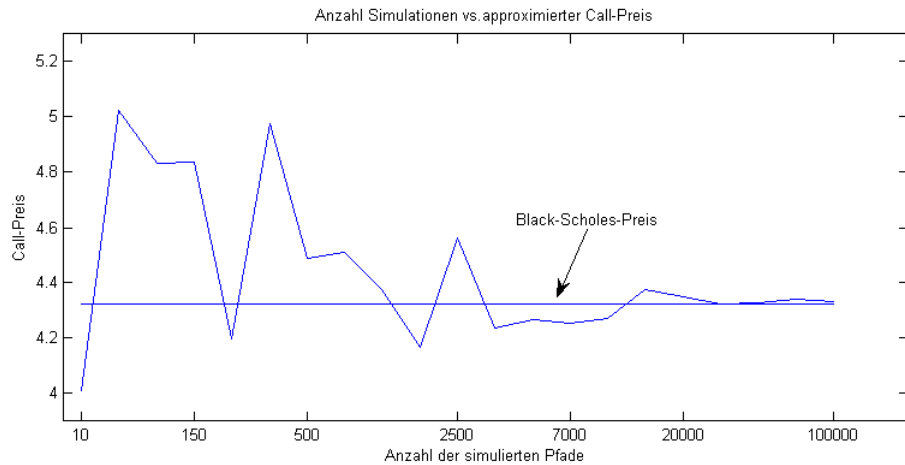


Abbildung 2: approximierter Call-Preis (mit Strikepreis $K=28$, Anfangswert $S_0=30$, $T=1$ und Zinsrate $r=0.05$) vs. Anzahl simulierter Pfade (im Vergleich zum Black-Scholes-Call-Preis 4.32)

zu erweitern, besteht darin einen Jump Prozess hinzuzufügen, welcher Sprünge des underlying assets zulässt.

Ein reiner Jump Prozess ist ein Prozess, dessen Pfade nicht mehr stetig sind, sondern dessen Werte sich nur durch Sprünge ändern. In diesem Abschnitt wird die Simulation von „zusammengesetzten“ Jump Prozessen eingeführt.

Bezeichnungen 3.7.

Wir betrachten einen stochastischen Prozess der Form:

$$J(t, \omega) = \sum_{j=1}^{N(t, \omega)} Y_j(\omega).$$

Dieser Prozess enthält zwei Terme, in denen Zufälligkeit steckt. Der Prozess N entscheidet über das Auftreten (insbesondere über die Anzahl) von Sprüngen. Deshalb wollen wir N einen Zählprozess nennen. Die auftretenden Sprungzeiten wollen wir mit $\tau_1, \dots, \tau_{N(t, \omega)}$ bezeichnen. Der zweite Zufälligkeit erzeugende Term Y bestimmt die Sprunghöhen, wobei Y_i die Sprunghöhe des i -ten Sprungs bezeichnet.

Wollen wir dann einen Jump Prozess J simulieren, so können wir die Simulation in die des Zählprozesses und die Erzeugung von Sprunghöhen aufteilen. Bei der Erzeugung von Sprunghöhen sind im Normalfall Verteilungen gegeben, weshalb wir diese mit einer statischen Monte Carlo Methode erzeugen können. Die Simulation von Sprungzeiten ist nicht ganz so einfach. Um die obige Bezeichnung noch zu verdeutlichen, geben wir die folgende

Definition 3.8 (Zählprozess).

Ein Zählprozess $N = (N(t, \omega), t \geq 0)$ ist ein monoton wachsender Prozess, welcher Werte in \mathbb{N} annimmt. Wir können diesen Prozess benutzen, um die Anzahl von

Sprüngen, welche im Intervall $[0, T]$ auftreten, zu modellieren, d.h. wir setzen $N_t :=$ „Anzahl Sprünge bis zur Zeit t “.

Im Folgenden werden wir, um Ereignisse - also Sprünge - zu zählen, Poisson-Prozesse verwenden:

Definition 3.9 (Poisson Prozess).

Ein Zählprozess $(N_t)_t$ heißt Poisson-Prozess mit Intensität λ , falls:

- (i) $N_0 = 0$,
- (ii) N hat unabhängige Zuwächse, d.h. für beliebige, sich nicht überlappende Intervalle $[s_i, t_i]$, $i \geq 0$, $t_i > s_i \geq 0$ ist $(N_{t_i} - N_{s_i})_i$ eine unabhängige Familie von Zufallsvariablen,
- (iii) N hat stationäre Poisson(λt)-verteilte Zuwächse, d.h. für $k \geq 0$

gilt: $\mathbb{P}(N_{t+s} - N_s = k) = \mathbb{P}(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$.

Bemerkung 3.10.

Seien die Wartezeiten zwischen dem Auftreten zweier Ereignisse $(\tau_i)_{i \geq 1}$ unabhängige, exponential-verteilte Zufallsvariablen mit Parameter λ (d.h. Zufallsvariablen mit Dichte $f(x; \lambda) = \lambda e^{-\lambda x} \mathbf{1}_{(0, \infty)}(x)$), so ist

$$N_{(t)} := \sum_{i \geq 1} \mathbf{1}_{t \geq T_i} \text{ mit } T_i := \sum_{j=1}^i \tau_j$$

ein Poisson-Prozess mit Intensität λ . (Bew.: siehe Anhang)

Zu bemerken sei die wichtige Eigenschaft des Poisson-Prozesses, dass die Zeiten zwischen dem Auftreten von Ereignissen (also die Zufallsvariablen $\tau_{i+1} - \tau_i$) exponentialverteilt und somit „gedächtnislos“ sind, d.h. das Auftreten von Ereignissen in der Vergangenheit unabhängig vom Auftreten eines Ereignisses (in den nachfolgenden Anwendungen also vom Auftreten von Sprüngen) in der Zukunft ist. Die Exponentialverteilung ist in der Tat die einzige stetige Verteilung mit dieser für uns wichtigen Eigenschaft.

Im Folgenden wollen wir N mittels der „Jump Intensität“ modellieren. Wir beginnen mit der Frage, was genau eine Jump Intensität oder auch Frequenz $\lambda > 0$, welche entweder konstant, deterministisch oder auch selbst zufällig sein kann, ist. Heuristisch ausgedrückt, definiert dieser Prozess die Anzahl von Sprüngen per Zeiteinheit in einer „kleinen“ Umgebung jedes Punktes $t \in [0, T]$.

In dieser Arbeit werden wir zunächst nur auf den Fall der konstanten Intensität weiter eingehen, der das Prinzip der Modellierung mittels der Intensität ausreichend klar machen wird. Um Pfade $(N_t)_{0 \leq t \leq T}$ und die korrespondierenden Sprungzeiten $\tau_1, \dots, \tau_{N_T}$ zu simulieren, werden wir die sogenannte Methode der „Countdown Simulation“ (nach [1]) verwenden.

Wir werden dazu annehmen, dass es sich bei N um einen Poisson-Prozess handelt, laut Definition also um einen Zählprozess mit unabhängigen und stationären Zuwächsen. Daraus folgt weiterhin, dass die Zeit zwischen zwei Sprüngen exponential verteilt ist zum Parameter λ . Wir können einen Zählprozess N daher wie folgt simulieren:

Algorithmus 3.11 (Countdown Simulation).

1. Setze $T_0 = 0, i = 0$
2. Setze $i = i + 1$
3. Erzeuge $U \sim \mathcal{U}[0, 1]$
4. Setze $\tau_i = -\lambda^{-1} \ln(U)$ und $T_i = T_{i-1} + \tau_i$
5. Wenn $T_i \leq T$, dann gehe zu Schritt 2
6. Output: Sprungzeiten T_1, \dots, T_{i-1} und das Poisson Sample $N_T = i - 1$

Bemerkung (zu Schritt 4). Ist U gleichverteilt auf $[0, 1]$, so ist $-\lambda^{-1} \ln(U)$ exponential verteilt zum Parameter λ , denn die Verteilungsfunktion der Exponentialverteilung lautet $F(x) = 1 - e^{-\lambda x}$, für $x \geq 0$ und daraus folgt mit Lemma 2.1, dass für F^{-1} , die (Pseudo-)Inverse der Verteilung der Exponentialverteilung, gilt $F^{-1}(U) = -\lambda^{-1} \ln(1 - U) \stackrel{\text{in Verteilung}}{=} -\lambda^{-1} \ln(U) \sim \text{Exp}(\lambda)$. Dass es sich wirklich um die (Pseudo-)Inverse handelt, zeigt $F^{-1}(F(x)) = -\lambda^{-1} \ln(1 - (1 - e^{-\lambda x})) = x$, für $x \geq 0$.

3.5 Mixed Jump Prozesse

Nach der Einführung der reinen Jump Prozesse wollen wir nun wie oben erwähnt Prozesse betrachten, deren Zufälligkeit sich aus einer Diffusionskomponente (einer zeitstetigen Brownschen Bewegung) und einem reinen Jump Prozess zusammensetzen. Wir betrachten im Folgenden also einen Prozess der Form

$$dX_t = \underbrace{\mu(t, X_{t-}) dt}_{\text{Drift}} + \underbrace{\sigma(t, X_{t-}) dW_t}_{\text{Volatilität}} + \underbrace{\eta(t, X_{t-})}_{\text{Einfluss der Sprünge}} dJ_t, \quad X_{t_0} = x,$$

wobei J ein zusammengesetzter Jump Prozess $J_t = \sum_{j=1}^{N_t} Y_j$ ist, $Y_j \stackrel{i.i.d.}{\sim} f_Y$ die zufälligen Sprunghöhen angibt und $\lambda(t, x)$ als Bestandteil des Poisson-Prozesses N die Jump Intensitätsfunktion bezeichnet. X_{t-} ist definiert als $\lim_{s \uparrow t} X_s$ und es gilt $X_{t-} = X_t$ für alle t , in denen kein Sprung stattfindet.

Wir wollen annehmen, dass $\eta(t, X_{t-})$ stetig in beiden Komponenten ist, sodass wir das Integral bezüglich „ dJ_t “ als Riemann-Stieltjes-Integral auffassen können, da ein Poisson-Prozess beschränkte Variation besitzt.

Wir nehmen im Folgenden weiter reguläre Koeffizienten an, um die Existenz und Eindeutigkeit einer Lösung der obigen stochastischen Differentialgleichung zu gewährleisten. Des Weiteren sei λ nach oben hin beschränkt.

Das Euler-Diskretisierungsverfahren liefert uns für den obigen Prozess dann folgende Approximation:

$$X_{i+1} = X_i + \mu(t_i, X_i) \Delta t + \sigma(t_i, X_i) \sqrt{\Delta t} \cdot Z_i + \eta(t_i, X_i) \cdot \sum_{j=N(t_i)+1}^{N(t_{i+1})} Y_j,$$

wobei Z_i natürlich wieder eine standardnormalverteilte Zufallsvariable ist und wir $\sum_{j=n+1}^n Y_j = 0 \forall n$ setzen.

Um gemischte Jump Prozesse zu simulieren, kommen wir nochmal auf die schon bekannte „Countdown Simulation“ 3.11 zurück. Diese wollen wir auf Prozesse mit

Sprüngen erweitern, welche sowohl von der Brownschen Bewegung, den Sprunghöhen als auch der Zustandsvariablen X selbst abhängen können. Im folgenden Beispiel wollen wir annehmen, dass die Intensitätsfunktion $\lambda(t, X_t)$ von der Zeitvariablen t und dem Prozess X_t abhängt. Nach [1] kann gezeigt werden, dass für jedes Paar von aufeinanderfolgenden Sprungzeiten τ_i und τ_{i+1} das Integral $\int_{\tau_i}^{\tau_{i+1}} \lambda(s, X_s) ds$ über die Intensitätsfunktion, gegeben τ_i , exponentialverteilt mit Parameter 1 ist. Im Spezialfall, dass $\lambda(t, X_t) = \lambda$ für alle t , sehen wir dies leicht ein, denn es gilt dann:

$$\int_{\tau_i}^{\tau_{i+1}} \lambda(s) ds = \int_{\tau_i}^{\tau_{i+1}} \lambda ds = \underbrace{(\tau_{i+1} - \tau_i)}_{\sim \text{Exp}(\lambda)} \lambda \stackrel{(*)}{\sim} \text{Exp}(\lambda/\lambda) = \text{Exp}(1).$$

(*) folgt aus der Tatsache, dass für alle γ und für alle $\text{Exp}(\lambda)$ -verteilte Zufallsvariablen X gilt, dass γX $\text{Exp}(\lambda/\gamma)$ -verteilt ist (siehe z.B. [3]). Für den Simulationsalgorithmus orientieren wir uns an der Idee aus [1] und gehen dann wie folgt vor:

Algorithmus 3.12 (Countdown Methode für Mixed Jump Prozesse).

1. Setze $X_0 = x, \tau_0 = 0, i = k = 0$ und wähle ein festes Δ_t
2. Erzeuge $e_k \sim \text{Exp}(1)$ unabhängig von e_{k-1} , falls dieses existiert
3. Setze $i = i + 1$; falls $i\Delta_t \geq T$, gehe direkt zu Schritt 9
4. Erzeuge $Z_i \sim \mathcal{N}(0, 1)$ unabhängig von allen vorherigen Samples
5. Setze $X_i = X_{i-1} + \mu(t_{i-1}, X_{i-1})\Delta_t + \sigma(t_{i-1}, X_{i-1})\sqrt{\Delta_t} \cdot Z_i$
6. Falls $\sum_{j=\tau_k/\Delta_t}^{i-1} \lambda(j\Delta_t, X_j)\Delta_t < e_k$, gehe zu Schritt 3
7. Setze $k = k + 1, \tau_k = i\Delta_t$
8. Erzeuge ein Sample der Sprunghöhe J_k , setze $X_i = X_i + J_k$ und gehe zu Schritt 2
9. Output: Sprungzeiten τ_1, \dots, τ_k und der Pfad aus Samples X_1, \dots, X_{i-1}

Um das Prinzip dieses Algorithmus zu verdeutlichen, machen wir noch ein paar Anmerkungen:

Nachdem wir in Schritt 2 ein Sample des Integrals über den Intensitätsprozess simuliert haben, testen wir in Schritt 6, ob wir - gegeben den letzten Sprungzeitpunkt τ_k - zum aktuellen Zeitpunkt $i\Delta_t$ den in Schritt 2 simulierten Wert des Integrals über die Intensität „schon erreicht“ haben. Dies würde nämlich bedeuten, dass wir „den Zeitpunkt des Sprungs erreicht haben“. Ist der Wert des approximierten Integrals ($\sum_{j=\tau_k/\Delta_t}^{i-1} \lambda(j\Delta_t, X_j)\Delta_t$) jedoch noch kleiner als der simulierte Wert, bei dessen Erreichen wir eine Sprung durchführen, so gehen wir einen Zeitschritt weiter, ohne dass ein Sprung stattfindet.

Zur Veranschaulichung eines gemischten Jump Prozesses betrachten wir noch einmal die geometrische Brownsche Bewegung, zu welcher wir nun einen Sprungsterm hinzufügen wollen. Bei der Darstellung orientieren wir uns an [6]. Wir nehmen an, dass die Sprungzeiten τ_1, \dots, τ_m in einem gegebenen Intervall bekannt sind. Der Sprungprozess $\sum_{j=1}^m Y_j$ soll folglich nur an den Sprungstellen Auswirkung auf die geometrische Brownsche Bewegung haben. Wir definieren $S_{t-} := \lim_{s \uparrow t} S_s$,

d.h. $S_{\tau_i^-}$ symbolisiert dann den Wert des Underlying „genau vor dem Sprung in τ_i “.

Wir betrachten nun eine geometrische Brownsche Bewegung mit Sprung-Term, gegeben durch die stochastische Differentialgleichung

$$dS_t = S_{t-} \mu dt + S_{t-} \sigma dW_t + S_{t-} dJ_t$$

und sehen, dass dS_t vom Wert des underlying assets genau vor dem möglichen Sprungzeitpunkt abhängt. Weiterhin gilt für die absolute Sprunghöhe:

$$S_{\tau_j} - S_{\tau_j^-} = S_{\tau_j^-} [J_{\tau_j} - J_{\tau_j^-}] = S_{\tau_j^-} Y_j, \text{ d.h.}$$

$$dJ_t = d \sum_{j=1}^{N_t} Y_j = \begin{cases} Y_j, & \text{falls } t = \tau_j \text{ für ein } j \leq m \\ 0, & \text{sonst} \end{cases}$$

gibt hier die relative Änderung von J_t im Intervall $[t, t+dt)$ an. Wir können zeigen, dass für den „Sprungterm“ folgendes gilt:

$$S_{t-} dJ_t = \sum_{s \in (t, t+dt]: \Delta S_s \neq 0} \Delta S_s,$$

wobei ΔS_s die „absolute Sprunghöhe im Zeitpunkt s “ bezeichne. Diese Tatsache stimmt mit unserer Anschauung, dass wir zum gegebenen stetigen Prozess an den Sprungstellen die absoluten Sprunghöhen hinzuaddieren, überein. Kommen wir zum Beweis dieser Erkenntnis:

Beweis.

Um zu verdeutlichen, dass es sich um die relative Änderung bzgl. t handelt, schreiben wir im Beweis d_t anstatt von d . Es gilt dann:

$$S_{t-} d_t J_t = S_{t-} d_t \sum_{j=1}^{N_t} Y_j = d_t \sum_{j=1}^{N_t} S_{\tau_j^-} \cdot Y_j,$$

weil $d_t \sum_{j=1}^{N_t} Y_j$ nur an den Sprungstellen Werte ungleich 0 annimmt.

Definieren wir $\Delta S_s := S_s - S_{s^-}$, so folgt weiter

$$\begin{aligned} d_t \sum_{j=1}^{N_t} S_{\tau_j^-} \cdot Y_j &= \sum_{j=1}^{N_t+dt} S_{\tau_j^-} \cdot Y_j - \sum_{j=1}^{N_t} S_{\tau_j^-} \cdot Y_j = \sum_{j=N_t+1}^{N_t+dt} S_{\tau_j^-} \cdot Y_j \\ &= \sum_{j=N_t+1}^{N_t+dt} S_{\tau_j} - S_{\tau_j^-} = \sum_{s \in (t, t+dt]: \Delta S_s \neq 0} \Delta S_s \end{aligned}$$

□

Wir bemerken noch, dass der Wert von S_t immer nichtnegativ bleibt, solange wir $Y_j \geq -1$ voraussetzen, was wir deshalb tun wollen. Nach diesem kleinen Einschub kommen wir nun zur Betrachtung der geometrischen Brownschen Bewegung mit

Jump-Term zurück.

Wie wir oben gesehen haben, gilt

$$S_{\tau_j} = S_{\tau_j^-} \cdot (1 + Y_j). \quad (3.2)$$

Wir erhalten folglich eine Darstellung der Werte des Prozesses zu den Sprungzeiten τ_i , $1 \leq i \leq m$, vorausgesetzt wir kennen den Wert des Prozesses „kurz vor dem Sprung“. Zwischen den Sprungzeiten folgt der Prozess aber den Dynamics einer gewöhnlichen geometrischen Brownschen Bewegung (da der Sprung-Term in der SDE wie gesagt nur an den Sprungstellen von 0 verschiedene Werte annimmt). Per Induktion können wir dann zeigen, dass die stochastische Differentialgleichung außerhalb der Sprungstellen die Lösung

$$S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t} \cdot \prod_{j=1}^{N_t} (1 + Y_j)$$

besitzt.

Beweis. [durch Induktion]

Wir betrachten für $n=1$ das Intervall $[0, \tau_1)$: Es gilt dann trivialerweise für alle $t \in [0, \tau_1)$ $S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t}$, da noch kein Sprung stattgefunden hat.

Für den Induktionsschritt nehmen wir an, dass die Behauptung für ein festes n gilt, also für alle $t \in [0, \tau_1) \cup (\tau_1, \tau_2) \cup \dots \cup (\tau_{n-1}, \tau_n)$.

Für $n \mapsto n + 1$ sei also $t \in (\tau_n, \tau_{n+1})$. Es gilt dann

$$\begin{aligned} S_t &= \underbrace{S_{\tau_n}}_{=S_{\tau_n^-}(1+Y_n) \text{ nach 3.2}} e^{(\mu - \sigma^2/2)(t - \tau_n) + \sigma(W_t - W_{\tau_n})} \quad (\text{geom. Brownsche Bew. mit Startwert } S_{\tau_n}) \\ &\stackrel{\text{Ind.schr.}}{=} S_0 \left(e^{(\mu - \sigma^2/2)\tau_n^- + \sigma W_{\tau_n^-}} \cdot \prod_{j=1}^{\overbrace{N_{\tau_n^-}}^{=n-1}} (1 + Y_j) \right) \cdot (1 + Y_n) \cdot e^{(\mu - \sigma^2/2)(t - \tau_n) + \sigma(W_t - W_{\tau_n})}. \end{aligned}$$

Mit der Stetigkeit von t und W_t , also mit $\tau_n^- = \tau_n$ und $W_{\tau_n^-} = W_{\tau_n}$ folgt weiter

$$\begin{aligned} S_t &= S_0 e^{(\mu - \sigma^2/2)\tau_n + \sigma W_{\tau_n}} \cdot \prod_{j=1}^{\overbrace{N_{\tau_n}}^{=n}} (1 + Y_j) \cdot e^{(\mu - \sigma^2/2)(t - \tau_n) + \sigma(W_t - W_{\tau_n})} \\ &= S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t} \cdot \prod_{j=1}^{N_t} (1 + Y_j) \end{aligned}$$

□

Insgesamt erhalten wir also die Lösung

$$S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t} \cdot \prod_{j=1}^{N_t} (1 + Y_j) \quad \forall t.$$

In der Abbildung 3 sehen wir auch anschaulich, dass ein gemischter Jump Prozess zwischen den Sprungzeiten nichts anderes als ein zeitstetiger Diffusionsprozess ist.

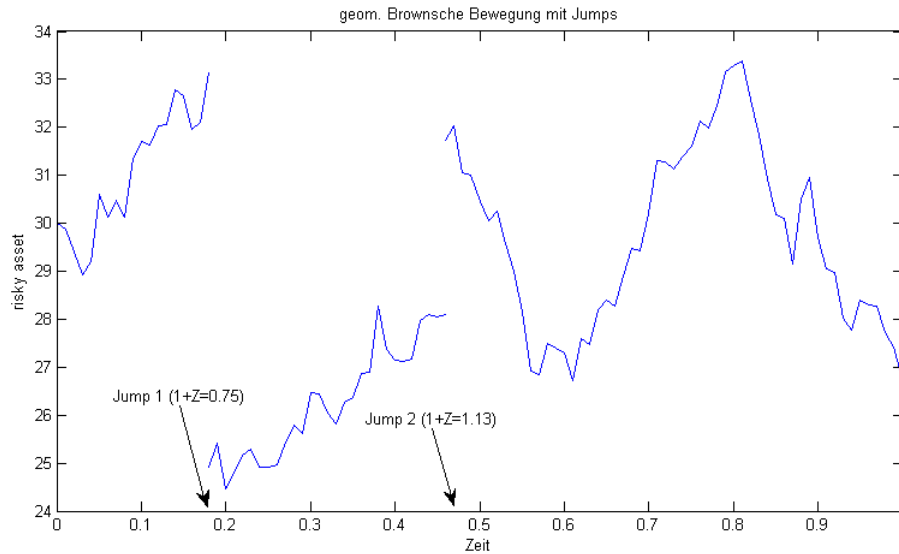


Abbildung 3: geometrische Brownsche Bewegung mit Jumps, wobei die Verteilung der Y_i gegeben ist durch $\ln(1+Y_i) =: \ln(1+Z) \sim \mathcal{N}(e, 0.5^2)$, $r=0.05$, $\sigma = 0.2$ und $\lambda = 3$ sei

An den simulierten Sprungzeiten multiplizieren wir den Wert des Prozesses $S(\tau_j^-)$ „kurz vor dem Sprung“ mit $(1+Y_i)$, um den Sprung durchzuführen. Bis zum nächsten Sprungzeitpunkt, welcher wie im oben behandelten Fall auch zufällig und in Abhängigkeit vom Zeitpunkt t und vom Prozess $X(t)$ gewählt werden kann, handelt es sich dann wieder um einen zeitstetigen Diffusionsprozess, in unserem Fall um eine geometrische Brownsche Bewegung. Wir können diese Vorgehensweise mit dem obigen Algorithmus „Countdown Methode für Mixed Jump Prozesse“ simulieren, falls wir J_k im obigen Algorithmus durch $S_{\tau_k} - Y_k$ simulieren, wobei $Y_k \geq -1$ den k -ten simulierten Wert einer Zufallsvariablen mit gegebener Verteilung, die die relative Sprunghöhe beschreiben kann, bezeichne.

Nach dieser Einführung in dynamische Monte Carlo Methoden und die Simulation von Pfaden von stochastischen Prozessen, wollen wir diese Methoden im nächsten Abschnitt anwenden.

4 Bewertung Asiatischer Optionen

Im folgenden Abschnitt wollen wir dynamische Monte Carlo Methoden nutzen, um Asiatische Optionen zu bewerten. Wir kommen dabei zum Black-Scholes Modell zurück. Bei Asiatischen Optionen handelt es sich nun nicht mehr um einfache Plain-Vanilla Optionen (also Optionen mit den typischen Eigenschaften wie Europäische Put-/Call-Optionen), sondern um sogenannte exotische Optionen. Die Asiatische Option ist eine pfadabhängige Option, bei der der Payoff vom durchschnittlichen Preis des Underlyings während der Laufzeit der Option abhängt. Dabei können wir verschiedene Typen von Asiatischen Optionen betrachten. Zunächst können wir die Option sowohl unter zeitstetiger Beobachtung als auch unter diskreter Beobachtung betrachten. Im zeitstetigen Fall wird der Mittelwert über den Beobachtungszeitraum als Integral berechnet, wohingegen wir den Mittelwert unter diskreter Beobachtung als Summe bestimmen. Darüberhinaus können Asiatische Optionen vom europäischen sowie vom amerikanischen Typ sein. Wir werden wie bisher jedoch nur auf den europäischen Typ eingehen. In dem Fall, dass wir die Option bzgl. des underlying assets S auf dem Intervall $[0, T]$ und mit Strike-Preis K bei zeitstetiger Beobachtung bewerten, beläuft sich der Payoff dieser Asiatischen Option im Zeitpunkt T auf

$$\left(\frac{1}{T} \int_0^T S_t dt - K\right)^+.$$

Dadurch, dass wir uns nun von gewöhnlichen Optionen abgewandt haben, bei denen wir im Black-Scholes Modell eine konkrete analytische Formel zur Berechnung von Preisen (Black-Scholes Formel) gegeben haben, gelangen wir nun zu Anwendungen der Monte Carlo Methoden, welche in der Realität auch verwendet werden. Denn sind mehr oder weniger „einfache“ Formeln zur Bewertung von Optionen gegeben, so benötigt man natürlich keine numerischen Verfahren, um Preise zu schätzen. Dies haben wir bisher gemacht, um die Verfahren einzuführen und zu veranschaulichen.

Bei Asiatischen Optionen ist die Bewertung mithilfe von Monte Carlo Methoden folglich besonders interessant, da es sich hier tatsächlich um die Berechnung von höher-dimensionalen Integralen handelt. Will man beispielsweise eine Asiatische Option über eine Laufzeit von einem Jahr bei diskreter täglicher Beobachtung (Handelstage) bewerten, so ist es notwendig, über 250 voneinander unabhängige Zufallsvariablen zu schätzen.

Wir bemerken, dass wir den Mittelwert bislang immer als arithmetisches Mittel berechnet haben. Bei arithmetischen Asiatischen Optionen gibt es keine Ergebnisse in geschlossener Form. Im Falle der zeitstetigen Beobachtung wurden aber bereits analytische Ausdrücke für die Laplace-Transformierte in der Laufzeitkomponente einer Asiatischen Call-Option hergeleitet (siehe [10] bzw. Geman und Yor (1993)). Für die Bewertung von Asiatischen Optionen, besonders bei diskreter Beobachtung, besitzt die Bewertung mit Monte Carlo Methoden in der Realität folglich eine große Bedeutung. Bevor wir mit der Simulation beginnen, sei noch angemerkt, dass eine andere Form Asiatischer Optionen darin besteht, die Berechnung des Mittelwerts mithilfe der geometrischen Summe durchzuführen, d.h.

also als

$$\sqrt[n]{\prod_{i=1}^n S_i}.$$

Wird das underlying asset durch eine geometrische Brownsche Bewegung simuliert (Black-Scholes Modell), sei also

$$S_i = S_{i-1} e^{(r - \frac{\sigma^2}{2})\Delta t + \sigma\sqrt{\Delta t}Z_i}, 1 \leq i \leq n,$$

wobei Z_i wieder eine standardnormalverteilte Zufallsvariable sei, so ist die geometrische Summe der S_i ein Produkt von lognormal verteilten Zufallsvariablen, welches wiederum lognormal verteilt ist, denn:

$$\sqrt[n]{\prod_{i=1}^n S_i} = \exp(\ln((\prod_{i=1}^n S_i)^{\frac{1}{n}})) = \exp(\frac{1}{n} \sum_{i=1}^n \ln(S_i)), \quad (4.1)$$

wobei die $\ln(S_i)$ normalverteilte Zufallsvariablen sind. Folglich ist auch die Summe $\sum_{i=1}^n \ln(S_i)$ normalverteilt und mit Gleichung 4.1 ist $\sqrt[n]{\prod_{i=1}^n S_i}$ wiederum lognormalverteilt. Das heißt, zur Bewertung Asiatischer Optionen mit geometrischer Mittelwertberechnung benötigen wir nicht unbedingt Monte Carlo Methoden, da wir hier einfache Verteilungen für die Bewertung zu Grunde liegen haben.

Wir gehen deshalb im Folgenden von Asiatischen Optionen mit arithmetischer Mittelwertberechnung unter diskreter Beobachtung aus. Der Wert einer derartigen Option ergibt sich zum Zeitpunkt T (Laufzeitende) zu

$$(\frac{1}{n} \sum_{i=1}^n S_i - K)^+,$$

wobei n die Anzahl der Beobachtungspunkte im zu betrachtenden Intervall $[0, T]$ beschreibt. In diesem Fall erhalten wir durch die Aufsummierung von lognormalverteilten Zufallsvariablen einen Ausdruck, der analytisch nur sehr schwer auszudrücken ist. Wir benutzen eine „einfache“ bzw. „naive“ Monte Carlo Methode, um den Preis einer Asiatischen Option unter diskreter Beobachtung zu berechnen.

4.1 „Naive“ Bewertung

Im Folgenden werden wir uns nun ausführlich mit der Bewertung von Asiatischen Optionen unter diskreter Beobachtung (und mit arithmetischer Mittelwertberechnung) beschäftigen. Dabei befinden wir uns weiterhin im Black-Scholes Modell. Wir haben oben schon erwähnt, dass es dabei nötig ist, n-dimensionale Integrale zu berechnen, weshalb Preise in der Praxis oft mithilfe von Monte Carlo Methoden bestimmt werden. Betrachten wir den Wert der Option zum Zeitpunkt T (Laufzeitende), so ergibt sich dieser, wie gesagt, zu

$$(\frac{1}{n} \sum_{i=1}^n S_i - K)^+.$$

In der folgenden Anwendung wollen wir die Anzahl der Beobachtungspunkte mit der Anzahl der Diskretisierungspunkte bei der Simulation der geometrischen Brownschen Bewegung gleichsetzen. Wir summieren also alle simulierten Werte des underlying assets auf und teilen durch deren Anzahl, um dessen durchschnittlichen Wert zu erhalten.

Ist der Pfad des underlying assets $(S_i)_i$ bekannt, so ergibt sich der Wert einer Asiatischen Option mit Laufzeit T , Strikepreis K , Startwert des Underlyings S_0 sowie einem risikolosen Zinssatz r im Zeitpunkt $t = 0$ zu

$$V_0 = e^{-rT} \max\left\{\frac{1}{n} \sum_{i=1}^n S_i - K, 0\right\}.$$

Gehen wir nun aber wieder davon aus, dass wir uns im Zeitpunkt $t=0$ befinden und den Pfad der geometrischen Brownschen Bewegung (des underlying assets) nicht kennen. Der Preis einer Asiatischen Call-Option ist der folgende:

$$C = e^{-rT} \mathbb{E}\left[\max\left\{\frac{1}{n} \sum_{i=1}^n S_i - K, 0\right\}\right], \quad (4.2)$$

wobei die S_i jeweils von einer standardnormalverteilten Zufallsvariable Z_i abhängen. Schreiben wir den obigen Ausdruck also als Integral, so sehen wir, dass es sich tatsächlich um ein Mehrfachintegral der Dimension n handelt:

$$C = e^{-rT} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \max\left\{\frac{1}{n} \sum_{i=1}^n S_i(z_i) - K, 0\right\} g(z_n) \cdots g(z_1) dz_n \cdots dz_1,$$

wobei g die Dichtefunktion der Standardnormalverteilung sei.

Um den obigen Erwartungswert 4.2 auf die einfachste bzw. „naive“ Weise zu schätzen, simulieren wir wie gewohnt n Punkte der geometrischen Brownschen Bewegung, summieren die Werte auf und teilen durch deren Anzahl n . Anschließend subtrahieren wir den Strikepreis und betrachten das Maximum dieses Wertes und 0. Durch Abzinsen bekommen wir den Wert der Option zum Zeitpunkt 0. Indem wir dieses m mal durchführen und ebenfalls den Mittelwert berechnen, erhalten wir einen Schätzwert für den gesuchten Preis, nämlich

$$\hat{V}_0 = \frac{1}{m} \sum_{j=1}^m \left[e^{-rT} \max\left\{\frac{1}{n} \sum_{i=1}^n S_i^{(j)} - K, 0\right\} \right],$$

wobei $S_i^{(j)}$ den Wert des j -ten simulierten underlying assets im Zeitpunkt $i\Delta_t$ darstellt.

4.2 Varianzreduktionsmethoden

Wie wir zu Beginn des Kapitels „Einblick in statische Monte Carlo Methoden“ festgestellt haben, gilt für den Approximationsfehler ϵ bei der Monte Carlo-Integration zur Schätzung des Erwartungswerts der Zufallsvariablen X

$$\epsilon \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

Dies motiviert den Versuch, die Varianz σ^2 von X zu verkleinern, ohne dabei den Erwartungswert von X zu verändern, um die Präzision der Monte Carlo Schätzung zu erhöhen. Dazu werden wir im kommenden Abschnitt zwei Varianzreduktionsmethoden vorstellen, die die Varianz deutlich verringern können, ohne den Erwartungswert zu verändern. Wir orientieren uns dabei an [11].

4.2.1 Kontrollvariablen

Die erste Varianzreduktionsmethode, die wir vorstellen möchten, benutzt sogenannte „Kontrollfunktionen“. Nehmen wir an, dass wir den Erwartungswert μ_X einer Zufallsvariablen X mit Varianz $Var(X) = \sigma_X^2$ schätzen wollen.

Weiter nehmen wir an, dass wir in der selben Simulation, in der wir X schätzen, weitere Samples einer Zufallsvariable $Y = (Y^{(1)}, \dots, Y^{(n)})$, $n \in \mathbb{N}$ mit bekanntem Erwartungswert μ_Y und (möglicherweise unbekannter) Varianz $Var(Y) = \sigma_Y^2$ erhalten können. Die Idee ist nun, vom „naiven“ Monte Carlo Schätzer X die einfach zu handhabende Kontrollvariable Y abzuziehen und deren Erwartungswert wieder hinzuzufügen. Wir definieren also

$$\hat{X} = X - b(Y - \mu_Y) \quad (4.3)$$

für einen Kontrollfaktor $b \in \mathbb{R}$. \hat{X} besitzt den selben Erwartungswert wie X und die Varianz

$$Var(\hat{X}) = Var(X) + b^2 Var(Y) - 2b Cov(X, Y) = \sigma_X^2 + b^2 \sigma_Y^2 - 2b Cov(X, Y).$$

Das diesen Ausdruck minimierende b^* erhält man mit einfacher Analysis als

$$b^* = \frac{Cov(X, Y)}{\sigma_Y^2}.$$

Weiterhin wissen wir, dass die Korrelation von X und Y definiert ist durch $Cor(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$, was uns insgesamt zu folgender minimaler Varianz von \hat{X} führt:

$$\begin{aligned} Var(\hat{X}) &= \sigma_X^2 + b^{*2} \sigma_Y^2 - 2b^* Cov(X, Y) = \sigma_X^2 - \frac{Cov(X, Y)^2}{\sigma_Y^2} \\ &= \sigma_X^2 - Cor(X, Y)^2 \sigma_X^2 = \sigma_X^2 (1 - Cor(X, Y)^2). \end{aligned}$$

Wir sehen also, dass die Varianz verringert wird, falls die Korrelation von X und Y nicht gleich 0 ist, da diese nur Werte im Intervall $[-1, 1]$ annimmt.

Das Problem, welches in der Praxis auftaucht, ist, dass die Kovarianz von X und Y typischerweise nicht gegeben ist. Deshalb muss man auch diese schätzen. Um Unabhängigkeit weiter zu gewährleisten, führen wir zunächst einen sogenannten „pilot run“ durch, also einen kürzeren Testlauf, bei dem wir das b^* schätzen. Seien dazu $(X^{(i)}, Y^{(i)})_{1 \leq i \leq m}$ m unabhängige Realisierungen von (X, Y) dann erhalten wir die Stichprobenkovarianz von X und Y , also eine Schätzung $\overline{Cov(X, Y)}$ für die Kovarianz von X und Y durch

$$\overline{Cov(X, Y)} = \frac{\sum_{i=1}^m (X^{(i)} - \hat{X})(Y^{(i)} - \hat{Y})}{m - 1}$$

und die Stichprobenvarianz von Y , also eine Schätzung $\overline{Var(Y)}$ für die Varianz von Y durch

$$\overline{Var(Y)} = \frac{\sum_{i=1}^m (Y^{(i)} - \tilde{Y})^2}{m-1}.$$

Damit erhalten wir dann eine Schätzung \tilde{b}^* für b^* durch

$$\tilde{b}^* = \frac{\sum_{i=1}^m (X^{(i)} - \tilde{X})(Y^{(i)} - \tilde{Y})}{\sum_{i=1}^m (Y^{(i)} - \tilde{Y})^2}, \quad (4.4)$$

wobei \tilde{X} und \tilde{Y} die durch den Testlauf geschätzten Mittelwerte seien. Wir erhalten dann den varianzreduzierten Schätzer

$$\hat{X} = X - \tilde{b}^*(Y - \mu_Y).$$

Bei der Monte Carlo Schätzung ersetzen wir selbstverständlich X und Y durch $\bar{X} = \frac{1}{n} \sum_{i=1}^n X^{(i)}$ bzw. $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y^{(i)}$.

Im folgenden Abschnitt wollen wir das gerade erklärte Verfahren für die Bewertung einer Asiatischen Option nutzen, wobei wir uns dazu an [12] orientieren wollen. Das Ziel ist es, den „naiven“ Schätzer $\frac{1}{m} \sum_{j=1}^m [e^{-rT} \max\{\frac{1}{n} \sum_{i=1}^n S_i^{(j)} - K, 0\}]$ zu verbessern. Wir definieren dazu die Zufallsvariable X als

$$X := e^{-rT} \max\{\frac{1}{n} \sum_{i=1}^n S_i - K, 0\}.$$

Im einleitenden Abschnitt oben haben wir bereits gesehen, dass wir neben arithmetischen Asiatischen Optionen auch Asiatische Optionen bzgl. des geometrischen Mittels betrachten können. Für diese existieren Verteilungen der Summe, weshalb wir Erwartungswerte berechnen können. Die erwartete Nähe von geometrischem und arithmetischem Mittel wollen wir ausnutzen und die Kontrollvariable

$$Y := e^{-rT} \max\{\sqrt[n]{\prod_{i=1}^n S_i} - K, 0\}$$

definieren. Wir wollen hier sowohl für die Schätzung des arithmetischen Mittels als auch für die des geometrischen Mittels die gleichen Simulationsergebnisse benutzen, weshalb die beiden Zufallsvariablen möglicherweise korreliert sind.

In Gleichung 4.1 haben wir gesehen, dass, wenn wir das underlying asset durch eine geometrische Brownsche Bewegung beschreiben, die geometrische Summe der beobachteten Werte S_i des Underlying die folgende Form hat:

$$P_n := \sqrt[n]{\prod_{i=1}^n S_i} = \exp\left(\frac{1}{n} \sum_{i=1}^n \ln(S_i)\right).$$

Weiterhin wissen wir, dass die S_i für alle $1 \leq i \leq n$ genauso verteilt sind wie $S_0 e^{(r - \frac{\sigma^2}{2})i\Delta_t + \sigma W_{i\Delta_t}}$. Daraus folgt weiter, dass P_n verteilt ist wie

$$\exp\left(\frac{1}{n} \sum_{i=1}^n \ln(S_0) \left[\left(r - \frac{\sigma^2}{2}\right) i \Delta_t + \sigma W_{i\Delta_t} \right]\right) = S_0 \exp\left(\frac{1}{n} \sum_{i=1}^n R(i)\right)$$

mit $R(i) := (r - \frac{\sigma^2}{2})i\Delta_t + \sigma W_{i\Delta_t} \sim \mathcal{N}\left(\left(r - \frac{\sigma^2}{2}\right)i\Delta_t, \sigma^2 i\Delta_t\right)$. Wir definieren weiter $\bar{R} := \frac{1}{n} \sum_{i=1}^n R(i)$ und berechnen unter Verwendung der Summenformel $\sum_{i=1}^n i = \frac{(n+1)n}{2}$

und der Definition $\Delta_t = \frac{T}{n}$ den Erwartungswert

$$\begin{aligned} \mathbb{E}[\bar{R}] &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n R(i)\right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[R(i)] = \frac{1}{n} \sum_{i=1}^n \left(r - \frac{\sigma^2}{2}\right) i \Delta_t = \frac{1}{n} \left(r - \frac{\sigma^2}{2}\right) \sum_{i=1}^n i \Delta_t \\ &= \left(r - \frac{\sigma^2}{2}\right) \frac{n+1}{2} \Delta_t = \left(r - \frac{\sigma^2}{2}\right) \frac{n+1}{2} \frac{T}{n} =: \tilde{\mu}T. \end{aligned}$$

Schreiben wir \bar{R} um zu

$$\begin{aligned} \bar{R} &= \frac{1}{n} \sum_{i=1}^n R(i) = \frac{1}{n} [nR(1) + (n-1)(R(2) - R(1)) + (n-2)(R(3) - R(2)) \\ &\quad + \dots + (R(n) - R(n-1))], \end{aligned}$$

so folgt mit der aus der Unabhängigkeit der Zuwächse der Brownschen Bewegung folgenden Unabhängigkeit der Zuwächse $R(i+1) - R(i)$ und unter Verwendung der Summenformel $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$ für die Varianz von \bar{R} dann

$$\begin{aligned} \text{Var}(\bar{R}) &= \frac{1}{n^2} [n^2 \text{Var}(R(1)) + (n-1)^2 \text{Var}(R(2) - R(1)) + \dots + \text{Var}(R(n) - R(n-1))] \\ &= \frac{1}{n^2} [n^2 \text{Var}\left(\left(r - \frac{\sigma^2}{2}\right)\Delta_t + \sigma W_{\Delta_t}\right) + (n-1)^2 \text{Var}\left(\left(r - \frac{\sigma^2}{2}\right)\Delta_t + \sigma(W_{2\Delta_t} - W_{\Delta_t})\right) \\ &\quad + \dots + \text{Var}\left(\left(r - \frac{\sigma^2}{2}\right)\Delta_t + \sigma(W_{n\Delta_t} - W_{(n-1)\Delta_t})\right)] \\ &= \frac{1}{n^2} [n^2 \sigma^2 \Delta_t + (n-1)^2 \sigma^2 \Delta_t + \dots + \sigma^2 \Delta_t] = \frac{1}{n^2} \sigma^2 \Delta_t \sum_{i=1}^n i^2 \\ &= \frac{1}{n^2} \sigma^2 \Delta_t \frac{n(n+1)(2n+1)}{6} = \frac{\sigma^2(n+1)(2n+1)}{6n} \Delta_t \\ &= \frac{\sigma^2(n+1)(2n+1)}{6n} \frac{T}{n} =: \tilde{\sigma}^2 T, \end{aligned}$$

d.h. \bar{R} ist $\mathcal{N}(\tilde{\mu}T, \tilde{\sigma}^2 T)$ verteilt. Daraus folgt für unsere Kontrollvariable Y dann

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[e^{-rT} \max\{S_0 e^{\bar{R}} - K, 0\}] \\ &= \mathbb{E}[e^{-rT} \max\{e^{(-r + \tilde{\mu} + \frac{\tilde{\sigma}^2}{2})T} S_0 e^{\bar{R} + (r - \tilde{\mu} - \frac{\tilde{\sigma}^2}{2})T} - K, 0\}], \end{aligned}$$

wobei $\bar{R} + (r - \tilde{\mu} - \frac{\tilde{\sigma}^2}{2})T \sim \mathcal{N}((r - \frac{\tilde{\sigma}^2}{2})T, \tilde{\sigma}^2 T)$. Wir sehen, dass wir an dieser Stelle die Black-Scholes Formel anwenden können, wenn wir in der üblichen Formel

$$BS(S_0, K, r, T, \sigma) = \mathbb{E}[e^{-rT} \max\{S_0 e^R - K, 0\}],$$

wobei $R \sim \mathcal{N}((r - \frac{\sigma^2}{2})T, \sigma^2 T)$, S_0 durch $S_0 e^{(-r + \tilde{\mu} + \frac{\tilde{\sigma}^2}{2})T}$ und σ durch $\tilde{\sigma}$ ersetzen. Wir kennen nun also den Erwartungswert von unser Kontrollvariablen Y , weshalb wir nun die neue varianzreduzierte Schätzvariable \hat{X} aus 4.3 definieren können als

$$\hat{X} = X - b^*(Y - \mathbb{E}[Y]) \text{ mit } b^* = \frac{Cov(X, Y)}{Var(Y)}.$$

Den Erwartungswert $\mathbb{E}[Y]$ können wir wie gezeigt mit der Black-Scholes Formel berechnen. Die Varianz von Y sowie die Kovarianz von X und Y werden wir - wie bei der Schätzung von \tilde{b}^* in Gleichung 4.4 erklärt - schätzen. Um diese Anwendung klarer zu machen, wollen wir einen Pseudo-Code angeben, mit welchem wir in diesem Modell den Preis einer Asiatischen Call-Option mithilfe der Kontrollvariablen $Y = e^{-rT} \max\{(\prod_{j=1}^n S_j)^{1/n} - K, 0\}$, welche wie oben den "zufälligen" Wert einer geometrischen Asiatischen Option beschreibt, berechnen können.

Algorithmus 4.1 (Preisberechnung einer Asiatischen Call-Option mit Kontrollvariable).

1. Input: m (# simulierte Pfade), n (# Beobachtungs-/Diskretisierungspunkte), T, S_0, r, σ, K, m' (# simulierte Pfade für Pilot-Run ($< m$))
2. Berechnung des Black-Scholes-Preises der geom. Asiatischen Option (= $\mathbb{E}[Y]$):
 - (a) $\tilde{\mu} = (r - \frac{\sigma^2}{2}) \frac{n+1}{2n}$
 $\tilde{\sigma}^2 = \frac{(n+1)(2n+1)\sigma^2}{6n^2}$
 - (b) Setze $BS_g = BS(S_0 e^{(-r + \tilde{\mu} + \frac{\tilde{\sigma}^2}{2})T}, K, r, T, \tilde{\sigma})$
3. Pilot-Run:
 - (a) Für $i=1:m'$
 - i. Erzeuge einen diskreten Pfad der geom. Brownschen Bewegung $(S_j)_{1 \leq j \leq n}^{(i)}$
 - ii. Berechne den „Wert des Pfades“ einer arithm. Asiatischen Option
 $X'^{(i)} := V_{0a}(S^{(i)}) = e^{-rT} \max\{\frac{1}{n} \sum_{j=1}^n S_j^{(i)}, 0\}$
 - iii. Berechne den „Wert des Pfades“ einer geom. Asiatischen Option
 $Y'^{(i)} := V_{0g}(S^{(i)}) = e^{-rT} \max\{(\prod_{j=1}^n S_j^{(i)})^{1/n}, 0\}$
 - (b) Berechne den Durchschnittswert der arithm. asiat. Option $\bar{X}' = \frac{1}{m'} \sum_{i=1}^{m'} X'^{(i)}$
 - (c) und den der geom. Asiatischen Option $\bar{Y}' = \frac{1}{m'} \sum_{i=1}^{m'} Y'^{(i)}$
 - (d) Berechne die Stichprobenvarianz des Wertes der geom. Asiatischen Option $Var(Y') = \frac{\sum_{i=1}^{m'} (Y'^{(i)} - \bar{Y}')^2}{m'-1}$
 - (e) und die Stichprobenkovarianz von $X' = (X'^{(1)}, \dots, X'^{(m')})$ und $Y' = (Y'^{(1)}, \dots, Y'^{(m')})$ $Cov(X', Y') = \frac{\sum_{i=1}^{m'} (Y'^{(i)} - \bar{Y}')(X'^{(i)} - \bar{X}')}{m'-1}$

- (f) setze $b^* = \frac{\text{Cov}(X', Y')}{\text{Var}(Y')}$
4. eigentliche Simulation:
- (a) Für $i=1:m$
- i. Erzeuge einen diskreten Pfad der geom. Brownschen Bewegung $(S_j)_{1 \leq j \leq n}^{(i)}$
 - ii. Berechne den „Wert des Pfades“ einer arithm. summierten Asiatischen Option $X^{(i)} := V_{0a}(S^{(i)}) = e^{-rT} \max\{\frac{1}{n} \sum_{j=1}^n S_j^{(i)}, 0\}$
 - iii. Berechne den „Wert des Pfades“ einer geom. summierten Asiatischen Option $Y^{(i)} := V_{0g}(S^{(i)}) = e^{-rT} \max\{\sqrt[n]{\prod_{j=1}^n S_j^{(i)}}, 0\}$
- (b) Berechne den Durchschnittswert der arithm. asiat. Option $\bar{X} = \frac{1}{m} \sum_{i=1}^m X^{(i)}$
- (c) und den der geom. Asiatischen Option $\bar{Y} = \frac{1}{m} \sum_{i=1}^m Y^{(i)}$
5. Output: Call-Preis der Asiatischen Option $C = \bar{X} - \beta^*(\bar{Y} - BS_g)$

Eine weiter Varianzreduktionsmethode, die wir im Folgenden kennenlernen wollen, ist das Importance Sampling.

4.2.2 Importance Sampling

Die Idee, die hinter der Importance Sampling Methode steckt, ist, durch einen Maßwechsel gewisse „wichtige“ Werte bei der Berechnung des Mittelwerts bei der einfachen Monte Carlo Methode mehr zu gewichten, um dadurch die Varianz zu verringern. Wir werden uns bei der Einführung dieser Methode wieder an [11] orientieren. Sei nun X ein m -dimensionaler Zufallsvektor, $f(x)$ die zugehörige m -dimensionale Dichtefunktion und $h(x)$ eine Funktion, sodass der Erwartungswert

$$\theta := \mathbb{E}_f[h(X)] = \int h(x)f(x)dx$$

existiere.

Sei $g(x)$ eine weitere Dichtefunktion, so dass der Träger von g Obermenge des Trägers von hf ist, also $\text{supp}(g) \supseteq \text{supp}(hf)$. Dann gilt

$$\theta = \mathbb{E}_f[h(X)] = \int_{x \in \text{supp}(g)} \frac{h(x)f(x)}{g(x)} g(x) dx = \mathbb{E}_g\left[\frac{h(X)f(X)}{g(X)}\right]. \quad (4.5)$$

\mathbb{E}_g bzw. \mathbb{E}_f bedeuten, dass die Erwartungswerte über eine bzgl. der Dichte g bzw. der Dichte f erzeugte Zufallsvariable gebildet werden. Wenden wir dies auf den „naiven“ Monte Carlo Schätzer $\hat{\theta}$ an, so gilt also

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n h(X_f^{(i)}) = \frac{1}{n} \sum_{i=1}^n \frac{f(X_g^{(i)})h(X_g^{(i)})}{g(X_g^{(i)})},$$

wobei die $X^{(i)}$ für alle i unabhängig identisch verteilt seien. In der ersten Summe werden sie allerdings bzgl. der Dichte f (deshalb X_f), in der zweiten Summe bzgl.

der Dichte g (deshalb X_g) erzeugt.

Für die Varianz von $\hat{\theta}$ folgt

$$\text{Var}(\hat{\theta}) \stackrel{X_g^{(i)} \text{ i.i.d.}}{=} \frac{1}{n} \sum_{i=1}^n \text{Var}\left(\frac{f(X_g^{(i)})h(X_g^{(i)})}{g(X_g^{(i)})}\right),$$

denn aus der Unabhängigkeit der $X_g^{(i)}$ folgt auch die Unabhängigkeit der $\frac{f(X_g^{(i)})h(X_g^{(i)})}{g(X_g^{(i)})}$

(vorausgesetzt die Funktionen sind messbar). Da $h(x)$ in der Anwendung gewöhnlich dem Wert eines Finanzderivats entspricht, wollen wir $h(x)$ als nichtnegativ voraussetzen. Es folgt dann $h(x) > 0$ für alle $x \in \text{supp}(hf)$.

Wir sehen, dass die Varianz von θ bzw. $\hat{\theta}$ minimiert wird, wenn wir $g(x)$ als Vielfaches von $h(x)f(x)$ wählen. Dazu benötigen wir das hf zu einer Dichte normierende Integral $\int_{\text{supp}(hf)} h(s)f(s)ds$. Dann könnten wir schreiben

$$g(x) = \frac{h(x)f(x)}{\int_{\text{supp}(hf)} h(s)f(s)ds}$$

und die Varianz von $\hat{\theta}$ wäre 0. Das Problem an dieser Überlegung ist, dass das Integral nicht lösbar ist. Der Grund, warum wir überhaupt die Monte Carlo Simulation durchführen, ist ja genau die mangelnde Berechenbarkeit eben dieses Integrals. In der Praxis ist diese ideale Wahl von g also nicht möglich. Immerhin wissen wir jetzt aber, dass wir g von möglichst ähnlicher Form wie hf wählen sollten.

Wir gehen deshalb nun davon aus, dass wir die Suche nach einer geeigneten Dichtefunktion $g(x)$ auf eine Klasse von Funktionen $g_\beta(x)$ einschränken können. β stellt hier den Parameter dar, für den wir im Folgenden einen optimalen Wert finden wollen. Optimal bedeutet hierbei natürlich die Varianz $\text{Var}(\hat{\theta})$ minimierend. Betrachten wir diese Varianz deshalb genauer. Es gilt

$$\begin{aligned} \text{Var}\left(\frac{h(X_{g_\beta})f(X_{g_\beta})}{g_\beta(X_{g_\beta})}\right) &= \mathbb{E}_{g_\beta}\left[\left(\frac{h(X)f(X)}{g_\beta(X)}\right)^2\right] - \mathbb{E}_{g_\beta}\left[\frac{h(X)f(X)}{g_\beta(X)}\right]^2 \\ &< \max_{x \in \text{supp}(g_\beta)} \left\{ \frac{h(x)f(x)}{g_\beta(x)} \right\} \mathbb{E}_{g_\beta}\left[\frac{h(X)f(X)}{g_\beta(X)}\right] - \theta_\beta^2 \\ &= \theta_\beta \left(\max_{x \in \text{supp}(g_\beta)} \left\{ \frac{h(x)f(x)}{g_\beta(x)} \right\} - \theta_\beta \right) =: \theta_\beta (M(\beta) - \theta_\beta), \end{aligned}$$

wobei θ_β hier klarmachen soll, dass θ noch vom Parameter β abhängt. Wir erhalten eine obere Grenze für die zu minimierende Varianz. Um das für diesen Fall optimale β zu finden, reicht es nun also den Ausdruck

$$M(\beta) = \max_{x \in \text{supp}(g_\beta)} \left\{ \frac{h(x)f(x)}{g_\beta(x)} \right\} \quad (4.6)$$

bzgl. β zu minimieren. Dies liegt an der Tatsache, dass für alle β gilt

$$M(\beta) \geq \mathbb{E}_{g_\beta}\left[\frac{h(X)f(X)}{g_\beta(X)}\right] = \theta_\beta$$

und es somit äquivalent ist den Ausdruck $\theta_\beta(M(\beta) - \theta_\beta)$ und den Ausdruck $M(\beta)$ zu minimieren.

Mithilfe dieses β bzw. der Dichte g_β können wir dann den Erwartungswert auf der rechten Seite der Gleichung 4.5 wie gewohnt als Monte Carlo Schätzwert berechnen.

Im Folgenden wollen wir diese Methode für die Bewertung einer Asiatischen Option nutzen. Sei das underlying asset wieder eine geometrische Brownsche Bewegung definiert durch

$$S_i = S_{i-1}e^{(r-\frac{\sigma^2}{2})\Delta_t + \sigma\sqrt{\Delta_t}Z_i} = S_0e^{(r-\frac{\sigma^2}{2})i\Delta_t + \sigma\sqrt{\Delta_t}\sum_{j=1}^i Z_j} \quad \text{mit } Z_i \stackrel{i.i.d.}{\sim} \mathcal{N}(0,1).$$

Dann gilt für den Call-Preis C einer Asiatischen Option

$$C = e^{-rT}\mathbb{E}_{Z\sim\mathcal{N}(0,I)}\max\left\{\frac{1}{n}\sum_{i=1}^n S_i - K, 0\right\} = \mathbb{E}_{Z\sim\mathcal{N}(0,I)}[V_0(Z)],$$

mit $Z = (Z_1, \dots, Z_n)^T$, $V_0(Z) := e^{-rT}\max\left\{\frac{1}{n}\sum_{i=1}^n S_i - K, 0\right\}$, wobei V_0 in dem Sinne von Z abhängt, als dass die S_i jeweils von Z_i abhängig sind. Wir schreiben aber statt $S_i(Z_i)$ im Folgenden weiter nur S_i und sind uns darüber bewusst, dass V_0 auf diese Weise indirekt von Z abhängt. Die Kovarianzmatrix I in der mehrdimensionalen Normalverteilung $\mathcal{N}(0, I)$ sei natürlich die Einheitsmatrix.

Wir gehen nun davon aus, dass unsere Importance Sampling Verteilung von der Form $\mathcal{N}(\beta, I)$ sei mit Parameter $\beta = (\beta_1, \dots, \beta_n)^T$, wir also einen Maßwechsel in Form einer Verschiebung des Erwartungswerts der Normalverteilung durchführen. Um die obigen Bezeichnungen beizubehalten, sei also $f(x)$ die Dichtefunktion der n-dimensionalen Standardnormalverteilung $\mathcal{N}(0, I)$, also $f(x) = \frac{1}{\sqrt{2\pi}^n}e^{-\frac{1}{2}x^T x}$, und $g_\beta(x)$ sei die Dichtefunktion der n-dimensionalen Normalverteilung $\mathcal{N}(\beta, I)$, also $g_\beta(x) = \frac{1}{\sqrt{2\pi}^n}e^{-\frac{1}{2}(x-\beta)^T(x-\beta)}$. Der Quotient $\frac{f(x)}{g_\beta(x)}$ ergibt sich dann als

$$\frac{f(x)}{g_\beta(x)} = \frac{e^{-\frac{1}{2}x^T x}}{e^{-\frac{1}{2}(x-\beta)^T(x-\beta)}} = e^{\frac{1}{2}\beta^T \beta - \beta^T x}.$$

Damit folgt mit Gleichung 4.5 für den Preis C der Asiatischen Option

$$C = \mathbb{E}_{Z\sim\mathcal{N}(0,I)}[V_0(Z)] = \mathbb{E}_{Z\sim\mathcal{N}(\beta,I)}[V_0(Z)e^{\frac{\beta^T \beta}{2} - \beta^T Z}].$$

Wir sehen, dass wir hier einen Maßwechsel durchführen, den wir bei der Simulation natürlich beachten müssen. Haben wir bei der Simulation der Pfade der geometrischen Brownschen Bewegung die Zuwächse ΔW_i des Wiener Prozesses W bislang für alle $1 \leq i \leq n$ durch $\Delta W_i = \sqrt{\Delta_t}Z_i \sim \mathcal{N}(0, \Delta_t)$ simuliert, so müssen wir diese nun durch die Zuwächse des Wiener Prozesses mit hinzugefügtem Erwartungswert $\sum_{j=1}^i \beta_j \sqrt{\Delta_t}$ ersetzen. Diese Zuwächse ergeben sich dann für alle i als $\Delta \tilde{W}_i = \sqrt{\Delta_t}(Z_i + \beta_i) \sim \mathcal{N}(\beta_i, \Delta_t)$.

Wie wir in der theoretischen Einführung gesehen haben, müssen wir nun den Ausdruck $M(\beta) = \max_{z \in \text{supp}(g_\beta)} \left\{ \frac{V_0(z)f(z)}{g_\beta(z)} \right\}$ aus Gleichung 4.6 bezüglich β minimieren.

Das β , für welches dieser Ausdruck minimal ist, wollen wir β^* nennen. Da wir innerhalb des Ausdrucks wiederum das Maximum über $\frac{V_0(z)f(z)}{g_\beta(z)}$ suchen und f und g_β Dichtefunktionen der Normalverteilung darstellen, also $f(z) > 0$, $g_\beta(z) > 0$ gilt, können wir den Fall, dass $V_0(z) = 0$ ist, vernachlässigen. Es reicht also $V_0(z) > 0$ zu betrachten und somit für dieses Problem $\sum_{i=1}^n S_i - K > 0$ vorauszusetzen.

Die notwendige Bedingung $\text{supp}(V_0 f) \subseteq \text{supp}(g_\beta)$ ist also auch sichergestellt. Betrachten wir nun wieder den Ausdruck

$$\begin{aligned} M(\beta) &= \max_{z \in \text{supp}(g_\beta)} \left\{ \frac{V_0(z)f(z)}{g_\beta(z)} \right\} = \max_{z \in \text{supp}(V_0 f)} \left\{ V_0(z) e^{\frac{\beta^T \beta}{2} - \beta^T z} \right\} \\ &= \max_{V_0(z) > 0} \left\{ e^{\ln(V_0(z)) + \frac{\beta^T \beta}{2} - \beta^T z} \right\}. \end{aligned}$$

Um den Ausdruck $e^{\ln(V_0(z)) + \frac{\beta^T \beta}{2} - \beta^T z}$ bzw. $\ln(V_0(z)) + \frac{\beta^T \beta}{2} - \beta^T z$ in z zu maximieren und den gesamten Ausdruck in β zu minimieren, ergeben sich analytisch die folgenden beiden Bedingungen:

$$\beta^* = \frac{\nabla V_0(z)}{V_0(z)} \text{ und } \beta^* = z.$$

Insgesamt suchen wir also das β^* mit

$$\beta^* = \frac{\nabla V_0(\beta^*)}{V_0(\beta^*)}. \quad (4.7)$$

Dieses wollen wir wie in [6] numerisch finden.

Dazu benötigen zunächst den Gradienten von $V_0 = e^{-rT} \max\left\{ \frac{1}{n} \sum_{i=1}^n S_i - K, 0 \right\}$. Da wir uns - wie oben erwähnt - beim Maximieren auf den Fall, dass $\frac{1}{n} \sum_{i=1}^n S_i - K > 0$ beschränken wollen, ist V_0 differenzierbar und für alle $1 \leq i \leq n$ gilt:

$$\begin{aligned} \frac{\partial V_0(\beta^*)}{\partial \beta_i^*} &= \frac{\partial(e^{-rT} (\frac{1}{n} \sum_{j=1}^n S_j - K))}{\partial \beta_i^*} = e^{-rT} \frac{1}{n} \sum_{j=1}^n \frac{\partial S_j}{\partial \beta_i^*} \\ &= e^{-rT} \frac{1}{n} \sum_{j=i}^n S_j \cdot \sigma \sqrt{\Delta_t}, \text{ da} \end{aligned} \quad (4.8)$$

$$\begin{aligned} \frac{\partial S_j}{\partial \beta_i^*} &= \frac{\partial(S_0 e^{(r - \frac{\sigma^2}{2})j\Delta_t + \sigma\sqrt{\Delta_t} \sum_{k=1}^j \beta_k^*})}{\partial \beta_i^*} \\ &= \begin{cases} (S_0 e^{(r - \frac{\sigma^2}{2})j\Delta_t + \sigma\sqrt{\Delta_t} \sum_{k=1}^j \beta_k^*}) \cdot \sigma \sqrt{\Delta_t} = S_j \cdot \sigma \sqrt{\Delta_t}, & \text{falls } i \leq j \\ 0, & \text{falls } i > j. \end{cases} \end{aligned}$$

Weiterhin bemerken wir, dass jeder Vektor von Samples der Standardnormalverteilung $Z = (Z_1, \dots, Z_n)$ eindeutig einen diskreten Pfad $S = (S_1, \dots, S_n)$ festlegt und damit auch den zugehörigen Payoff V_0 .

Nehmen wir an, wir kennen den Wert V_0 und setzen diesen gleich y , so können wir die β_i^* mit den Gleichungen 4.7 und 4.8 iterativ berechnen. Wir beginnen mit $\beta^*_1(y)$:

$$\begin{aligned}\beta^*_1(y) &= \frac{e^{-rT} \frac{1}{n} \sum_{j=1}^n S_j \cdot \sigma \sqrt{\Delta t}}{V_0} = \frac{\sigma \sqrt{\Delta t} (e^{-rT} (\frac{1}{n} \sum_{j=1}^n S_j - K + K))}{V_0} \\ &= \frac{\sigma \sqrt{\Delta t} (V_0 + e^{-rT} K)}{V_0} = \frac{\sigma \sqrt{\Delta t} (y + e^{-rT} K)}{y}\end{aligned}$$

Als nächstes können wir dann

$$S_1(y) = S_0 e^{(r - \frac{\sigma^2}{2})\Delta t + \sigma \sqrt{\Delta t} \beta^*_1(y)}$$

berechnen, daraus wiederum $\beta^*_2(y) = \beta^*_1(y) - \frac{e^{-rT} \frac{1}{n} S_1(y) \sigma \sqrt{\Delta t}}{y}$, denn

$$\begin{aligned}\beta^*_{i+1} - \beta^*_i &= \frac{e^{-rT} \frac{1}{n} \sum_{j=i+1}^n S_j \cdot \sigma \sqrt{\Delta t}}{y} - \frac{e^{-rT} \frac{1}{n} \sum_{j=i}^n S_j \cdot \sigma \sqrt{\Delta t}}{y} = \frac{e^{-rT} \frac{1}{n} (-S_i \cdot \sigma \sqrt{\Delta t})}{y} \\ \Rightarrow \beta^*_{i+1} &= \beta^*_i - \frac{e^{-rT} \frac{1}{n} S_i \cdot \sigma \sqrt{\Delta t}}{y}.\end{aligned}$$

Iterativ erhalten wir so die von y abhängigen $\beta^*_i(y)$ und $S_i(y)$. Anstatt das geeignete β^* zu finden, müssen wir nun also nur das y finden, welches den Payoff $e^{-rT} (\frac{1}{n} \sum_{i=1}^n S_i(y) - K) = V_0 = y$ liefert, also die Lösung der Gleichung

$$y = e^{-rT} \left(\frac{1}{n} \sum_{i=1}^n S_i(y) - K \right)$$

berechnen.

Nach [6] zeigen numerische Versuche, dass die Gleichung eine eindeutige Lösung besitzt, welche sich beispielsweise leicht mit dem Intervallschachtelungsverfahren approximieren lässt. Wir können die soeben kennengelernte Methode wieder in einem Pseudo-Code zusammenfassen:

Algorithmus 4.2 (Preisberechnung einer Asiatischen Call-Option mit Importance Sampling).

1. Input: m (# simulierte Pfade), n (# Beobachtungs-/Diskretisierungspunkte), T , S_0 , r , σ , K
2. Berechnung von y als Lösung der Gleichung $y = e^{-rT} (\frac{1}{n} \sum_{i=1}^n S_i(y) - K)$ (mit Intervallschachtelung):
 - (a) Berechne iterativ die $\beta_i(y)$ und $S_i(y)$ mit:

- i. $\beta_1(y) = \frac{1}{y}(\sigma\sqrt{\Delta t}(y - e^{-rT}K))$
 - ii. $S_1(y) = S_0e^{(r-\frac{\sigma^2}{2})\Delta t + \sigma\sqrt{\Delta t}\beta_1(y)}$
 - iii. Für $i=2:n$
 - A. $\beta_i(y) = \beta_{i-1}(y) - \frac{e^{-rT}\frac{1}{n}S_{i-1}(y)\sigma\sqrt{\Delta t}}{y}$
 - B. $S_i(y) = S_{i-1}e^{(r-\frac{\sigma^2}{2})\Delta t + \sigma\sqrt{\Delta t}\beta_i(y)}$
- (b) Definiere damit die Funktion $f(y) = e^{-rT}(\frac{1}{n}\sum_{i=1}^n S_i(y) - K) - y$ und führe das Intervallschachtelungsverfahren zur Nullstellensuche durch.
- (c) Output: Nullstelle y^* , welche folglich die Gleichung $y = e^{-rT}(\frac{1}{n}\sum_{i=1}^n S_i(y) - K)$ löst.
3. Berechne den Vektor $\beta(y^*)$ mithilfe des obigen Algorithmus (aus 2.(a))
 4. Berechnung des Call-Preises der Asiatischen Option bezüglich des geänderten Maßes (statt in der Simulation der Pfade der geometrischen Brownschen Bewegung $(S_i)_i$ die $Z_i \sim \mathcal{N}(0, I)$ zu simulieren, simulieren wir diese Z_i als zufällige Werte der Normalverteilung mit Erwartungswert $\beta \mathcal{N}(\beta, I)$):
 - (a) Für $j=1:m$
 - i. Simuliere einen diskreten Pfad der geometrischen Brownschen Bewegung (beachte den Maßwechsel):
Für $i=1:n$
 - $\tilde{Z}_i = \beta_i + Z_i$ mit $Z_i \sim \mathcal{N}(0, 1)$
 - $W_i = \sqrt{\Delta t}(\tilde{Z}_i)$
 - $S_i = S_{i-1}e^{(r-\frac{\sigma^2}{2})\Delta t + \sigma W_i}$
 - ii. Berechne den „Wert des Pfades“ unter dem neuen Maß:
 $\tilde{C}^{(j)} = e^{-rT} \max\{\frac{1}{n}\sum_{i=1}^n S_i - K, 0\}$
 - iii. Rechne den Wert um ins ursprüngliche Maß: $C^{(j)} = \tilde{C}^{(j)} e^{\frac{\beta^T \beta}{2} - \beta^T \tilde{Z}}$
- (b) Berechne den Call-Preis der Asiatischen Option als Mittelwert der einzelnen „Werte der Pfade“:

$$C = \frac{1}{m} \sum_{j=1}^m C^{(j)}$$

4.3 Berechnung von Preisen und Fehlerabschätzung

Zum Abschluss der Bewertung von Asiatischen Call-Optionen im Black-Scholes Modell wollen wir noch ein paar mithilfe der in diesem Kapitel erarbeiteten Algorithmen berechnete Preise angeben. Um die drei Verfahren vergleichen zu können, wollen wir noch einmal auf einen sinnvollen Fehlerbegriff eingehen. Da sich Preise Asiatischer Optionen nicht einfach mit der Black-Scholes Formel oder anderen analytischen Formeln berechnen lassen, können wir den Fehler der Monte Carlo Methoden nicht als Differenz zum „richtigen“ (Black-Scholes-) Preis angeben. Wir wollen Fehler deshalb über die zu beobachtende Stichprobenvarianz abschätzen. Entstehen beim Berechnen von Call-Preisen hohe Streuungen, so liegt es nahe,

dass die Methode ungenauer ist als eine Methode, bei welcher nur kleine Streuungen auftreten. Wir wollen im Folgenden deshalb für alle drei in diesem Abschnitt vorgestellten Verfahren jeweils 50 mal den gleichen Call-Preis berechnen, diese Preise mitteln und die Stichprobenvarianz der Preise als Vergleichskriterium betrachten. In unserem Beispiel seien $r=0.1$, $K=100$, $\sigma=0.4$, $T=1$, $n=250$ (Diskretisierungspunkte), $m=50000$ (Pfade) und $m'=12400$ (Pfade beim Pilot-Run der Kontrollvariablenmethode). Wir variieren den Startwert des underlying assets S_0 . Es ergeben sich dann die folgenden Call-Preise und Varianzen:

S_0	Naive Simulation		Kontrollvariable		Importance Sampling	
	Preis	Varianz	Preis	Varianz	Preis	Varianz
95	8.4014	$4.00 \cdot 10^{-3}$	8.3999	$0.011693 \cdot 10^{-3}$	8.3974	$0.33592 \cdot 10^{-3}$
100	11.1618	$4.30 \cdot 10^{-3}$	11.1606	$0.011568 \cdot 10^{-3}$	11.1656	$0.56883 \cdot 10^{-3}$
105	14.2994	$6.70 \cdot 10^{-3}$	14.3080	$0.022483 \cdot 10^{-3}$	14.3126	$0.63322 \cdot 10^{-3}$

Wir sehen, dass beide Varianzreduktionsmethoden erfolgreich sind und die Varianz deutlich verringern. Vergleicht man die beiden Varianzreduktionsmethoden miteinander, so sieht man, dass die Kontrollvariablenmethode noch deutlich genauere Ergebnisse liefert. Allerdings müssen wir dazu sagen, dass diese im Vergleich zum Importance Sampling das fünf- bis sechsfache der Rechenzeit benötigte.

4.4 Put-Call-Parität

Nehmen wir an, wir befinden uns weiter im Black-Scholes Modell. Von Europäischen Plain-Vanilla Optionen kennen wir bereits die Put-Call Parität, welche besagt, dass

$$C_t - P_t = S_t - Ke^{-r(T-t)} \quad \forall 0 \leq t \leq T,$$

wobei C_t der Call-Preis der Option mit Laufzeit T zum Zeitpunkt t und P_t der Put-Preis der Option mit Laufzeit T zum Zeitpunkt t sei. Verwenden wir diese Put-Call-Parität, so ist es leicht, mithilfe des Call-Preises, den wir mit den Monte Carlo Methoden schätzen, auch den Put-Preis einer Option zu berechnen. Um auch den Put-Preis Asiatischer Optionen berechnen zu können, leiten wir nun eine ähnliche Put-Call-Parität für arithmetische Asiatische Optionen bei diskreter Beobachtung her. Die Werte der Asiatischen Call- bzw. Put-Optionen im Zeitpunkt t ($= \Delta_t i$ für ein $0 \leq i \leq n$) ergeben sich bekannterweise als

$$C_t = e^{-r(T-t)} \mathbb{E}[\max\{\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} - K, 0\} | \mathcal{F}_t] \text{ bzw.}$$

$$P_t = e^{-r(T-t)} \mathbb{E}[\max\{K - \frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j}, 0\} | \mathcal{F}_t].$$

Bei $(\mathcal{F}_t)_t$ handelt es sich wieder um die vom zugrundeliegenden Wiener Prozess erzeugte Filtration, was hier bedeutet, dass wir die Werte des underlying assets $(S_s)_s$ für alle $s \leq t$ kennen. Es ergibt sich dann die

Satz 4.3 (Put-Call-Parität für Asiatische Optionen).

Für Asiatische Optionen gilt:

$$C_t - P_t = \frac{S_t}{n - (t/\Delta_t)} \left[\frac{1 - e^{-r(T-t)}}{1 - e^{-r\Delta_t}} - e^{-r(T-t)} K \right] \text{ für alle } t = \Delta_t i \text{ mit } 0 \leq i \leq n.$$

Beweis. Mit einer einfachen Fallunterscheidung für die Fälle, dass

$\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} < K$ und $\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} > K$ sehen wir, dass

$$\begin{aligned} & \mathbb{E}[\max\{\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} - K, 0\} | \mathcal{F}_t] - \mathbb{E}[\max\{K - \frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j}, 0\} | \mathcal{F}_t] \\ &= \mathbb{E}[\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} - K | \mathcal{F}_t]. \end{aligned}$$

Weiter folgt mit Gleichung 2.1 von Seite 19, dass für alle $1 \leq j \leq n-i$

$$\mathbb{E}[S_{t+\Delta_t j} | \mathcal{F}_t] = \mathbb{E}[S_t | \mathcal{F}_t] e^{r\Delta_t j} = S_t e^{r\Delta_t j}.$$

Damit folgt:

$$\begin{aligned} C_t - P_t &= e^{-r(T-t)} [\mathbb{E}[\max\{\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} - K, 0\} | \mathcal{F}_t] - \mathbb{E}[\max\{K - \frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j}, 0\} | \mathcal{F}_t]] \\ &= e^{-r(T-t)} \mathbb{E}[\frac{1}{n-i} \sum_{j=1}^{n-i} S_{t+\Delta_t j} - K | \mathcal{F}_t] \\ &\stackrel{\text{Lin.}}{=} e^{-r(T-t)} [\frac{1}{n-i} \sum_{j=1}^{n-i} \mathbb{E}[S_{t+\Delta_t j} | \mathcal{F}_t] - K] \\ &= e^{-r(T-t)} [\frac{1}{n-i} \sum_{j=1}^{n-i} S_t e^{r\Delta_t j} - K] = e^{-r(T-t)} [\frac{S_t}{n-i} \sum_{j=1}^{n-i} (e^{r\Delta_t})^j - K] \\ &\stackrel{(1)}{=} e^{-r(T-t)} [\frac{S_t}{n-i} (\frac{e^{r\Delta_t((n-i)+1)} - e^{r\Delta_t}}{e^{r\Delta_t} - 1}) - K] \\ &= \frac{S_t}{n-i} (\frac{e^{-rT+rt+r\Delta_t n-r\Delta_t i+r\Delta_t} - e^{-rT+rt+r\Delta_t}}{e^{r\Delta_t} - 1}) - e^{-r(T-t)} K \\ &\stackrel{(2)}{=} \frac{S_t}{n - (t/\Delta_t)} (\frac{e^{r\Delta_t} - e^{-r(T-t)+r\Delta_t}}{e^{r\Delta_t} - 1}) - e^{-r(T-t)} K \\ &\stackrel{(3)}{=} \frac{S_t}{n - (t/\Delta_t)} (\frac{1 - e^{-r(T-t)}}{1 - e^{-r\Delta_t}}) - e^{-r(T-t)} K \end{aligned}$$

Im Umformungsschritt (1) haben wir die geometrische Reihe für $x \neq 1$

$$\sum_{i=0}^m x^i = \frac{x^{(m+1)} - 1}{x - 1} \Rightarrow \sum_{i=1}^m x^i = \frac{x^{(m+1)} - x}{x - 1}$$

für $x = e^{r\Delta t} \neq 1$ benutzt. In Umformungsschritt (2), dass $t = \Delta t i$ sowie, dass $\Delta t n = T$ gilt. In (3) erweitern wir den Bruch

$$\frac{e^{r\Delta t} - e^{-r(T-t)+r\Delta t}}{e^{r\Delta t} - 1}$$

mit $e^{-r\Delta t}$. □

4.5 Das stochastische Volatilitätsmodell nach Heston

Bisher haben wir uns darauf beschränkt, Preise im Black-Scholes Modell zu berechnen. Dazu mussten wir immer eine geometrische Brownsche Bewegung simulieren, deren Pfade den Werteverlauf des underlying assets beschreiben sollen. Bei „Plain-Vanilla“ Optionen können wir den Call- bzw. Put-Preis mithilfe der berühmten Black-Scholes Formel berechnen. In der Realität können wir jedoch beobachten, dass Optionspreise, welche mit der Black-Scholes Formel berechnet werden, nicht genau mit den am Markt zu beobachtenden Preisen übereinstimmen. Dies liegt unter anderem daran, dass die Volatilität σ im Black-Scholes Modell als konstant angenommen wird. Am Markt ist aber zu beobachten, dass die implizite Volatilität, also die Volatilität, die durch empirische Optionspreise bestimmt wird, stark von der Laufzeit und dem Strike-Preis (bzw. der Moneyness) abhängt. Betrachtet man die Black-Scholes Formel als Funktion des Marktpreises von Optionen, welche die zugehörige implizite Volatilität angibt, (d.h. alle anderen Kennzahlen seien fest gewählt,) so kann man abhängig vom Marktpreis die passende implizite am Markt vorherrschende Volatilität bestimmen. Trägt man die implizite Volatilität in einem Diagramm abhängig vom Strike-Preis auf, so entsteht eine Art „smile“; d.h. die minimale Volatilität ist etwa beim „at-the-money“ Strike-Preis zu finden, wohingegen die Volatilität sowohl unter- als auch oberhalb davon ansteigt. Dieser sogenannte „Volatilitäts-Smile“ führte dazu, dass Modelle entwickelt wurden, in denen die Volatilität σ als von der Zeit abhängige, also deterministische, Funktionen angenommen wird. Um die Zufälligkeit weiter zu erhöhen, wurden dann sogenannte „stochastische Volatilitätsmodelle“ entwickelt, bei denen die Volatilität, wie der Name schon vermuten lässt, durch einen stochastischen Prozess beschrieben wird. Eines der bekanntesten stochastischen Volatilitätsmodelle ist das von Heston. Das underlying asset S wird dabei durch die folgenden stochastischen Differentialgleichungen beschrieben:

Definition 4.4 (Das stochastische Volatilitätsmodell nach Heston - nach [13]).

$$\begin{aligned} dS_t &= S_t(rdt + \sqrt{v_t}dW_t^1) \\ dv_t &= \zeta(\bar{v} - v_t)dt + \alpha\sqrt{v_t}dW_t^2 \end{aligned}$$

Dabei ist \bar{v} der Langzeit- bzw. Mittelwert der Volatilität, ζ der Mean-Reversion Faktor, d.h. der Faktor, der bestimmt, wie stark der Volatilitätsprozess von seinem Mittel- bzw. Langzeitwert „angezogen“ werden soll und α die Volatilität der Volatilität. Wir wollen nun wieder arithmetische Asiatische Optionen bewerten. Nehmen

wir dazu an, dass W_t^1 und W_t^2 Wiener Prozesse seien und $W_t = (W_t^1, W_t^2)^T \mathcal{N}(0, \Sigma)$ -verteilt sei mit Kovarianzmatrix

$$\Sigma = t \cdot \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \text{ mit } \rho \in [-1, 1].$$

Offensichtlich müssen wir zunächst wissen, wie genau wir mehrdimensionale normalverteilte Zufallsvariablen, insb. also die Korrelation von W_t^1 und W_t^2 , simulieren können.

Im allgemeinen Fall, also für $d \geq 2$ sei dazu $Z = (Z^{(1)}, \dots, Z^{(d)})^T$ ein Vektor aus unabhängigen standardnormalverteilten Zufallsvariablen $Z^{(i)}$ und Σ eine symmetrische und positiv definite $d \times d$ -Matrix. Aus der linearen Algebra wissen wir, dass nun eine invertierbare Matrix L existiert mit $\Sigma = LL^T$ (siehe bspw. auch [3]). Wir behaupten nun, dass $Y := LZ$ dann $\mathcal{N}(0, \Sigma)$ -verteilt ist.

Satz 4.5 (aus [3]).

Gegeben $d \geq 2$, ein d -dimensionaler Zufallsvektor $Z = (Z^{(1)}, \dots, Z^{(d)})^T$ mit unabhängigen, jeweils standard-normalverteilten Komponenten, eine invertierbare, reelle $d \times d$ -Matrix L sowie ein Vektor $\mu = (\mu_1, \dots, \mu_d)^T \in \mathbb{R}^d$, besitzt $LZ + \mu$ die Dichte

$$\phi_{\mu, \Sigma}(x) = \frac{1}{(2\pi)^{d/2} (\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right),$$

wobei $\Sigma = LL^T$ und folglich $\det \Sigma = (\det L)^2$ gilt.

Wählen wir $\mu = 0$, so liefert uns dieser Satz, dass $Y = LZ$ tatsächlich $\mathcal{N}(0, \Sigma)$ -verteilt ist. Aus der Numerik der linearen Algebra wissen wir, dass die Matrix L mithilfe der Cholesky-Zerlegung für symmetrische und positiv definite Matrizen als untere Dreiecksmatrix bestimmt werden kann. Ein Pseudo-Code aus [6] befindet sich im Anhang.

Im Fall, dass Y und Z zweidimensional sind, also $Y = (Y^{(1)}, Y^{(2)})^T$ und $Z = (Z^{(1)}, Z^{(2)})^T$, und die Kovarianzmatrix von Y gegeben ist durch

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

ergibt sich folgendes:

Wir wählen den Ansatz, dass L die Form

$$L = \begin{pmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{pmatrix}$$

besitzt, womit uns die Cholesky-Zerlegung liefert, dass

$$\Sigma = LL^T \Rightarrow \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} \\ 0 & l_{22} \end{pmatrix} = \begin{pmatrix} l_{11}^2 & l_{11}l_{21} \\ l_{11}l_{21} & l_{21}^2 + l_{22}^2 \end{pmatrix},$$

woraus sich wiederum für L leicht ergibt, dass

$$L = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1 - \rho^2} \end{pmatrix}.$$

Wollen wir nun Y simulieren, so reicht es uns wieder $Z^{(1)} \sim \mathcal{N}(0, 1)$ und $Z^{(2)} \sim \mathcal{N}(0, 1)$ zu simulieren, denn wir setzen dann

$$Y = (Y^{(1)}, Y^{(2)})^T = L(Z^{(1)}, Z^{(2)})^T = (\sigma_1 Z^{(1)}, \rho\sigma_2 Z^{(1)} + \sigma_2 \sqrt{1 - \rho^2} Z^{(2)})^T.$$

Kommen wir nun zu unserem Ausgangsproblem zurück, bei welchem wir Pfade des mehrdimensionalen Prozesses $W_t = (W_t^1, W_t^2)^T$, wobei W_t^1 und W_t^2 Wiener Prozesse seien und deren Kovarianzmatrix durch

$$\Sigma = t \cdot \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \text{ mit } \rho \in [-1, 1]$$

gegeben sei, simulieren wollen. Mit dem soeben gezeigten und $\sigma_1 = \sigma_2 = \sqrt{t}$ wissen wir nun, dass wir die Pfade der W_t^1 und W_t^2 simulieren können durch

$$W_t^1 = \sigma_1 Z^{(1)} = \sqrt{t} Z^{(1)} \text{ und } W_t^2 = \rho\sigma_2 Z^{(1)} + \sigma_2 \sqrt{1 - \rho^2} Z^{(2)} = \sqrt{t}(\rho Z^{(1)} + \sqrt{1 - \rho^2} Z^{(2)}),$$

wobei $Z^{(1)}$ und $Z^{(2)}$ standardnormalverteilte Zufallsvariablen seien.

Im nächsten Schritt „diskretisieren“ wir die beiden stochastischen Differentialgleichungen mit dem Euler-Maruyama Verfahren, um Pfade simulieren zu können. Die beiden das underlying asset beschreibenden Gleichungen lauten dann:

$$\begin{aligned} v_{i+1} &= v_i + \zeta(\bar{v} - v_i)\Delta_t + \alpha\sqrt{v_i}\sqrt{\Delta_t}(\rho Z_i^{(1)} + \sqrt{1 - \rho^2} Z_i^{(2)}) \\ S_{i+1} &= S_i + S_i(r\Delta_t + \sqrt{v_i}(\sqrt{\Delta_t} Z_i^{(1)})) \end{aligned}$$

für $1 \leq i \leq n$ und gegebene v_0 und S_0 . Damit erzeugen wir wieder m (diskrete) Pfade $(S_i^{(j)})_i, 1 \leq j \leq m$ des underlying assets, womit wir dann wie üblich den Preis einer Asiatischen Call-Option berechnen können über

$$C = \frac{1}{m} \sum_{j=1}^m (e^{-rT} \max\{\frac{1}{n} \sum_{i=1}^n S_i^{(j)} - K, 0\}).$$

Im Anhang befindet sich - neben den Codes für die Berechnung der Call-Preise der Asiatischen Optionen aus der obigen Tabelle - auch ein Matlab-Code, welcher Preise von Asiatischen Call-Optionen im Heston Modell berechnen kann.

5 Fazit

In dieser Arbeit haben wir gesehen, dass dynamische Monte Carlo Methoden ein wichtiges Hilfsmittel, vor allem zur Bewertung von pfadabhängigen Optionen darstellen. Dies haben wir am Beispiel der Asiatischen Optionen herausgestellt. Wie wir gesehen haben, besitzen Monte Carlo Methoden meist ein im Vergleich zu anderen Methoden relativ hohes Maß an Ungenauigkeit, weshalb wir Methoden kennengelernt haben, die die Streuung der berechneten Preise minimieren. Bei der Berechnung von höher-dimensionalen Integralen bzw. Erwartungswerten stellen sich Monte Carlo Methoden so als sinnvolle Möglichkeit zur Berechnung von Optionspreisen heraus. Der große Vorteil der Monte Carlo Methoden besteht darin, dass sie immer anwendbar sind, auch wenn die Dimension des zu berechnenden „Integrals“ sehr hoch ist, was u.a. bei Asiatischen Optionen der Fall ist.

Da in der Praxis meist - wahrscheinlich der Einfachheit halber - das Euler-Maruyama Diskretisierungsverfahren benutzt wird, haben wir uns in dieser Arbeit ebenfalls darauf beschränkt. Für weiterführende Arbeiten wäre es interessant zu schauen, wie groß der Vorteil bei der Konvergenzgeschwindigkeit ist, wenn wir stattdessen mit dem Milstein-Verfahren diskretisieren würden. Weiterhin sind wir im überwiegenden Teil der Arbeit davon ausgegangen, dass wir uns im Black-Scholes Modell befinden. Im letzten Teil haben wir jedoch gesehen, dass es möglich ist, die kennengelernten Verfahren auf komplexere Finanzmarktmodelle - wie sie in der Realität verwendet werden - zu erweitern. Am Ende des Kapitels, in der wir dynamische Monte Carlo Methoden eingeführt haben, sind wir beispielsweise auf Mixed Jump Prozesse eingegangen, welche neben der stochastischen Volatilität eine weitere Möglichkeit darstellen, das Black-Scholes Modell zu erweitern. Sicherlich ist es interessant auch in solchen Modellen Optionspreise mit Monte Carlo Methoden zu berechnen. Ein weiteres Beispiel für pfadabhängige Optionen, für welche Monte Carlo Methoden in der Praxis relevant sind, sind doppelte Barriereoptionen. Deshalb ist sicher auch die Anwendung der Verfahren auf diesen Optionstyp interessant.

6 Anhang

Markov-Ungleichung und Borel-Cantelli-Lemma

Nach [3] gilt der folgende :

Satz (Markov-Ungleichung). Sei X eine reelle Zufallsgröße auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist für jede monoton wachsende Funktion $g : [0, \infty) \rightarrow [0, \infty)$ und jedes $t > 0$ mit $g(t) > 0$:

$$\mathbb{P}(|X| \geq t) \leq g(t)^{-1} \int_{|X| \geq t} g(|X|) d\mathbb{P} \leq g(t)^{-1} \mathbb{E}[g(|X|)].$$

Im obigen Beweis nutzen wir die Ungleichung für $g(t) = t^2$, also als

$$\mathbb{P}(|X| \geq t) \leq t^{-2} \int_{|X| \geq t} X^2 d\mathbb{P} \leq t^{-2} \mathbb{E}[X^2]$$

Weiterhin gilt nach [3] auch das folgende

Lemma (von Borel-Cantelli). Sei $(A_n)_{n \geq 1}$ eine Folge von Ereignissen in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann gilt

$$\sum_{n \geq 1} \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0.$$

Sind die $A_n, n \geq 1$, außerdem unabhängig, gilt auch die Umkehrung oder, äquivalent hierzu:

$$\sum_{n \geq 1} \mathbb{P}(A_n) = \infty \Rightarrow \mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 1.$$

Dabei gilt

$$\limsup_{n \rightarrow \infty} A_n = \{\omega \in \Omega : \omega \text{ liegt in unendlich vielen } A_n\}$$

Beweis von Bemerkung 3.10:

Aussage: Seien die Wartezeiten zwischen dem Auftreten zweier Ereignisse $(\tau_i)_{i \geq 1}$ unabhängige, exponential-verteilte Zufallsvariablen mit Parameter λ , so ist

$$N_{(t)} := \sum_{i \geq 1} \mathbf{1}_{t \geq T_i} \text{ mit } T_i := \sum_{j=1}^i \tau_j$$

ein Poisson-Prozess mit Intensität λ .

Beweis. (i) $N_{(0)} = 0$, denn: $\mathbf{1}_{t \geq T_i} = 0$ für $t=0$, da $T_i \geq \tau_1 > 0 \forall i \geq 1$

(ii) Die Unabhängigkeit der Zuwächse folgt aus der „Gedächtnislosigkeit“ der Exponentialverteilung

(iii)

$$\begin{aligned}\mathbb{P}(N_{(t)} = k) &= \mathbb{P}(\tau_1 + \dots + \tau_k \leq t, \tau_1 + \dots + \tau_{k+1} > t) \\ &= \mathbb{P}(\tau_1 + \dots + \tau_k \leq t) \mathbb{P}(\tau_{k+1} > t - (\tau_1 + \dots + \tau_k) \mid \tau_1 + \dots + \tau_k \leq t) \\ &\stackrel{(*)}{=} \int_0^t \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x) dx \cdot \left(\int_{t-x}^{\infty} \lambda e^{-\lambda y} \mathbb{1}_{(0,\infty)}(y) dy \cdot \mathbb{1}_{x \leq t} \right) \\ &= \frac{\lambda^{k+1}}{(k-1)!} \int_0^t x^{k-1} e^{-\lambda x} dx \cdot \int_t^{\infty} e^{-\lambda(y-x)} \mathbb{1}_{x \leq t}(y) dy \\ &= \frac{\lambda^{k+1}}{(k-1)!} \int_0^t x^{k-1} e^{-\lambda x} \cdot \frac{1}{\lambda} e^{-\lambda(t-x)} dx \\ &= \frac{\lambda^k}{(k-1)!} \int_0^t x^{k-1} e^{-\lambda t} dx \\ &= \frac{\lambda^k}{(k-1)!} e^{-\lambda t} \frac{1}{k} (t^k - 0^k) \\ &= \frac{(t\lambda)^k e^{-\lambda t}}{k!} \\ &\Rightarrow N_{(t)} \sim \text{Poisson}(\lambda t)\end{aligned}$$

(*) folgt aus der Tatsache, dass die Summe von k unabhängig identisch exponential(λ)-verteilten Zufallsvariablen nach [3] Erlang($k;\lambda$)-verteilt ist, also die Dichte

$$\lambda^k \frac{x^{k-1}}{(k-1)!} e^{-\lambda x} \mathbb{1}_{(0,\infty)}(x)$$

besitzt, sowie daraus, dass τ_{k+1} $Exp(\lambda)$ -verteilt ist. □

Cholesky-Zerlegung für symmetrische und positiv definite Matrizen - nach [6]

Algorithmus.

1. Input: symmetrische und positiv definite $d \times d$ -Matrix Σ
2. Initialisiere L: $L = \text{zeros}(d,d)$
3. Für $j=1:d$
 - Für $i=j:d$
 - (a) $v_i = \Sigma_{ij}$
 - (b) Für $k=1:j-1$
$$v_i = v_i - L_{jk} L_{ik}$$
 - (c) $L_{ij} = v_i / \sqrt{v_j}$
4. Output: Untere Dreiecksmatrix L

Tabelle:

Approximierte Europäische Call-Preise in Abhängigkeit der Anzahl der simulierten Pfade und Approximationsfehler (Underlying asset: Geometrische Brownsche Bewegung)

# simulierter Pfade	approx.Call-Preis	abs.Fehler im Vgl.zum BS-Preis	relativer Fehler
10	4.0089	-0.3156	8,24 %
50	5.0241	0.6997	16,20 %
100	4.8344	0.5100	11,81 %
150	4.8350	0.5106	11,82 %
200	4.1980	-0.1264	2,93 %
250	4.9750	0.6505	15,06 %
500	4.4871	0.1627	3,77 %
750	4.5096	0.1852	4,29 %
1250	4.3766	0.0522	1,21 %
1750	4.1687	-0.1557	3,60 %
2500	4.5626	0.2382	5,51 %
3500	4.2363	-0.0881	2,04 %
5000	4.2694	-0.0550	1,27 %
7000	4.2546	-0.0698	1,62 %
10000	4.2734	-0.0510	1,18 %
15000	4.3745	0.0500	1,16 %
20000	4.3515	0.0270	0,63 %
25000	4.3231	-0.0013	0,03 %
35000	4.3272	0.0028	0,06 %
50000	4.3429	0.0185	0,43 %
100000	4.3343	0.0099	0,23 %

Matlab-Codes für die Berechnung von Call-Preisen Asiatischer Optionen

Naive Monte Carlo Schätzung des Call-Preises einer Asiatischen Option

```
% Dieses Programm simuliert m Pfade der geometrischen Brownschen
% Bewegung und berechnet in Abhängigkeit von m den Call-Preis
% einer Asiatischen Option als Mittelwert der einzelnen auf Basis der
% Pfade berechneten Preise (im BS-Modell)
```

```
m=50000 % Anzahl der simulierten Pfade
n=250 % Anzahl der Diskretisierungsintervalle
```

```
% Initialisierungen:
W=zeros(n+1,1); S=W; C_T=zeros(m,1); C=C_T;
```

```
S_0=100 % Startwert
T=1 % Endzeitpunkt
s=0.4 % Volatilität
K=100 % Strike-Preis
r=0.1 % risikoloser Zinssatz
dt=T/n; % Intervallbreite
```

```
for j=1:m
```

```

% Erzeuge Pfad der geom. Brownschen Bewegung:
W(1)=0;
S(1)=S_0;
for i=1:n
    % Sample random noise:
    W(i+1)=randn;
    % Berechne Wert des Underlying:
    S(i+1)=S(i)*exp((r-s^2/2)*dt+s*sqrt(dt)*W(i+1));
end
% Berechne "Wert des Pfades": Preis zum Zeitpunkt t_0=0 unter Kenntnis des
% Pfades auf [t_0,T]:
C_T(j)=max((sum(S)-S(1))/n-K,0); % Wert der asiatischen Calloption in T
C(j)=C_T(j)*exp(-r*T); % Wert der Calloption in t_0
end
y=mean(C) % approximierter Call-Preis (in Abh. von m)

```

Monte Carlo Schätzung des Call-Preises einer Asiatischen Option mit Kontrollvariable

Hauptprogramm:

```

% Dieses Programm berechnet den Preis einer arithm. asiatischen Option
% mit n Diskretisierungspunkten (=Beobachtungspunkte) mithilfe von Kontroll-
% variablen zur Varianzreduktion im Black-Scholes Modell. Dazu berechnet es
% den Black-Scholes Preis einer geometrischen asiatischen Option, welche
% als Kontrollvariable genutzt wird, führt einen Pilot-Run durch, um den
% optimalen Kontrollfaktor b* zu berechnen, berechnet die naiv geschätzten
% Werte der geometrischen asiatischen Option und der arithm. asiatischen
% Option, um zuletzt den Call-Preis mithilfe des varianzreduzierten
% Schätzers  $X_v = X - b*(Y - BSg)$  zu berechnen.

```

```

m=50000 % Anzahl der simulierten Pfade
n=250 % Anzahl der Diskretisierungsintervalle

```

```

S_0=100 % Startwert
T=1 % Endzeitpunkt
s=0.4 % Volatilität
K=100 % Strike-Preis
r=0.1 % risikoloser Zinssatz

```

```

dt=T/n; %Intervallbreite

```

```

mu2=(r-s^2/2)*(n+1)/(2*n) % \tilde{\mu}
s2=sqrt((n+1)*(2*n+1)*s^2/(6*n^2)) % \tilde{\sigma}

```

```

BSg=blsprice(S_0*exp((-r+mu2+s2^2/2)*T), K, r, T, s2) % Black-
Scholes-Preis der geometrischen asiatischen Call-Option

```



```

% Pilot-Run
mm=round(m/4); % Pilot-Run mit kleinerer Anzahl mm von berechneten Pfaden
X=zeros(mm,1); Y=X; % Initialisierungen
for i=1:mm
    S=gBB(n,S_0,T,s,r); % erzeuge einen Pfad der geom. Brownschen Bew.
    C_a=max(sum(S)/n-K,0); % "Wert des Pfades" (arithm.Call) in T
    X(i)=C_a*exp(-r*T); % abgezinster "Wert des Pfades"(arithm.Calloption)
    C_g=max(prod(nthroot(S,n))-K,0); %"Wert des Pfades" (geom.Call) in T
    Y(i)=C_g*exp(-r*T); % abgezinster "Wert des Pfades" (geom.Calloption)
end
XX=mean(X); % Durchschnittspreis ("naiver Schätzer") arithm. asiat. Option
YY=mean(Y); % Durchschnittspreis ("naiver Schätzer") geom. asiat. Option

Sum1=0; Sum2=0;
for i=1:mm
    Sum1=Sum1+(Y(i)-YY)^2; % geschätzte Varianz(X)*(mm-1)
    Sum2=Sum2+(X(i)-XX)*(Y(i)-YY); % geschätzte Kovarianz(X,Y)*(mm-1)
end
b=Sum2/Sum1 % optimaler Kontrollfaktor b*
% Ende: Pilot-Run
% eigentliche Simulation:
A=zeros(m,1); G=A; % Initialisierungen
for j=1:m
    S=gBB(n,S_0,T,s,r); % erzeuge einen Pfad der geom. Brownschen Bew.
    C_a=max(sum(S)/n-K,0); % "Wert des Pfades" (arithm.Call) in T
    A(j)=C_a*exp(-r*T); % abgezinster "Wert des Pfades"(arithm.Calloption)
    C_g=max(prod(nthroot(S,n))-K,0); %"Wert des Pfades" (geom.Call) in T
    G(j)=C_g*exp(-r*T); % abgezinster "Wert des Pfades" (geom.Calloption)
end
C=mean(A)-b*(mean(G)-BSg) % geschätzter Call-Preis d.arithm. asiat.Option

gBB:
% Dieses File simuliert einen Pfad der geometrischen Brownschen Bewegung
% Annahmen: t_0=0
function S=gBB(n,S_0,T,s,r)
% Initialisierungen:
W=zeros(n,1); S=W;

dt=T/n; % Intervallbreite

% Erzeuge Pfad der geom.Brownschen Bewegung:
W(1)=randn;
S(1)=S_0*exp((r-s^2/2)*dt+s*sqrt(dt)*W(1));
for i=2:n
    % Sample random noise:
    W(i)=randn;

```

```

    % Berechne Wert des Underlying:
    S(i)=S(i-1)*exp((r-s^2/2)*dt+s*sqrt(dt)*W(i));
end
end

```

Monte Carlo Schätzung des Call-Preises einer Asiatischen Option mit Importance Sampling

Hauptprogramm:

```

% Dieses Programm berechnet den Preis einer arithm. asiatischen Option mit
% n Diskretisierungspunkten (=Beobachtungspunkte) mithilfe von Importance
% Sampling zur Varianzreduktion im Black-Scholes Modell. Dazu löst es
% zunächst die Gleichung  $y = \exp(-r \cdot T) \cdot (1/n \cdot (\sum S_i) - K)$ , um dann mithilfe
% dieses y den Maßwechsel—"Verschiebungsvektor" beta auszurechnen.
% Anschließend werden m geometrische Brownsche Bewegungen simuliert
% (beachte den Maßwechsel: Statt die  $Z_i$  in der geometrischen Brownschen
% Bewegung standardnormalverteilt zu simulieren, simulieren wir diese
% normalverteilt mit E-Wert beta.
% Damit wird dann wie gewohnt der Wert jedes Pfades bzgl. des neuen Maßes
%  $P_{\{Z+\beta\}}$  berechnet und der zugehörige Wert des Calls C bezüglich des
% Ausgangsmaßes  $P_{\{Z\}}$  berechnet durch
%  $C (=E_{\{Z\}}(V_T)) = E_{\{Z+\beta\}}(V_T \cdot \exp(\beta \cdot \beta' / 2 - \beta \cdot Z'))$ . Die
% einzelnen m Werte werden wie gewohnt gemittelt, um den Preis zu erhalten.

```

```

m=50000 % Anzahl der simulierten Pfade
n=250 % Anzahl der Diskretisierungsintervalle

```

```

S_0=100 % Startwert
T=1 % Endzeitpunkt
s=0.4 % Volatilität
K=100 % Strike-Preis
r=0.1 % risikoloser Zinssatz

```

```

dt=T/n; % Intervallbreite
% Initialisierungen:
C_T_M=zeros(1,m); C_M=C_T_M; C=C_T_M;

```

```

% Berechnung von y als Lösung der Gleichung  $y = \exp(-r \cdot T) \cdot (1/n \cdot (\sum S_i) - K)$ :
y=BisectionIS(n,r,s,T,K,S_0);
% Berechnung des Vektors beta in Abhängigkeit von y:
beta=ImpSamp_beta(y,n,r,s,T,K,S_0,dt);
% Berechnung des Call Preises mit der geometrischen Brownschen Bewegung
% bzgl. des neuen Maßes:
for j=1:m
Z=zeros(1,n+1); S=Z; y=0;
% Erzeuge Pfad der geom.Brownschen Bewegung:

```

```

Z(1)=0;
S(1)=S_0;
for i=1:n
    % Simuliere normalverteilte ZV mit E-Wert beta(i):
    Z(i+1)=(beta(i)+randn);
    % Berechne Wert des Underlying im Zeitpunkt dt*i:
    S(i+1)=S(i)*exp((r-s^2/2)*dt+s*sqrt(dt)*Z(i+1));
end
% Berechne "Wert des Pfades": Preis zum Zeitpunkt t_0=0 unter Kenntnis des
% Pfades auf [t_0,T] (unter neuem Maß!):
C_T_M(j)=max((sum(S)-S(1))/n-K,0); % Wert der asiat. Calloption in T
C_M(j)=C_T_M(j)*exp(-r*T); % Wert der Calloption in t_0 (unter neuem Maß)
C(j)=C_M(j)*exp(beta*beta'/2-beta*Z(2:n+1)'); % Umrechnung zu
    gewohntem Maß
end
y=mean(C) % approximierter Call-Preis (in Abh. von m)

```

ImpSamp_beta:

```

% Dieses Programm berechnet in Abhängigkeit von y die Summe
% S:=1/n*(sum(S_i(y))) für das Importance Sampling, sowie den
% zugehörigen Vektor beta und gibt den Vektor beta aus.
function beta=ImpSamp_beta(y,n,r,s,T,K,S_0,dt)
bS=zeros(2,n);
bS(1,1)=(s*sqrt(dt)*(y+exp(-r*T)*K))/y; % beta_1(y)
bS(2,1)=S_0*exp((r-s^2/2)*dt+s*sqrt(dt)*bS(1,1)); % S_1(y)
for i=2:n
    bS(1,i)=bS(1,i-1)-(exp(-r*T)*1/n*s*sqrt(dt)*(bS(2,i-1)))/y; %
        beta_i(y)
    bS(2,i)=bS(2,i-1)*exp((r-s^2/2)*dt+s*sqrt(dt)*bS(1,i)); % S_i(y)
end
beta=bS(1,:); % Output: beta
end

```

Monte Carlo Schätzung des Call-Preises einer Asiatischen Option im Heston Modell

```

% Dieses File erzeugt einen zufälligen Pfad des underlying assets im
% HESTON-MODELL (mithilfe des Euler-Verfahrens) und berechnet den Monte
% Carlo Schätzpreis einer arithmetischen asiatischen Call-Option

```

```

m=50000 % Anzahl simulierte Pfade

```

```

n=250 % Anzahl der Diskretisierungsintervalle
r=0.05 % risikoloser Zinssatz
S_0=30 % Startwert
T=1/2 % Endzeitpunkt

```

```

K=28 % Strike-Preis

dt=T/n; % Intervallbreite

s1=0.3 % Volatilität des 1.Wienerprozesses
s2=0.3 % Volatilität des 2.Wienerprozesses
rho=0.2 % Korrelation der Wiener Prozesse
sigma=[s1^2,rho*s1*s2;rho*s1*s2,s2^2] % Kovarianzmatrix
zeta=2 % Mean-Reversion Faktor
vv=0.3 % Langzeitvolatilität des Underlying assets
alpha=0.2 % Volatilität der Volatilität des underlying assets

%Initialisierungen:
C_T=zeros(1,m); C=zeros(1,m);

for j=1:m
v=zeros(1,n+1);
S=zeros(1,n+1);
v(1)=0.25; % Anfangsvolatilität des Underlying assets
S(1)=S_0;

% Simulation der Pfade:
for i=1:n
    % Zweidimensionale Standardnormalverteilung Z(Z_1,Z_2)^T:
    Z=randn(2,1);
    % Simulation der beiden Wiener Prozesse W1 und W2:
    dW1=Z(1)*sqrt(dt);
    dW2=(rho*Z(1)+sqrt(1-rho^2)*Z(2))*sqrt(dt);
    % Berechnung der Pfade der Volatilität und des Underlying assets:
    v(i+1)=v(i)+zeta*(vv-v(i))*dt+alpha*v(i)*dW2;
    S(i+1)=S(i)+S(i)*(r*dt+v(i)*dW1);
end
% Berechne "Wert des Pfades": Preis zum Zeitpunkt t_0 unter Kenntnis des
% Pfades auf [t_0,T]:
C_T(j)=max((sum(S)-S(1))/n-K,0); % Wert der asiatischen Calloption in T
C(j)=C_T(j)*exp(-r*T); % Wert der Calloption in t_0
end
y=mean(C) % approximierter Call-Preis (in Abh. von m)

```

Literatur

- [1] G.FUSAI, A.RONCORONI: *Implementing Models in Quantitative Finance: Methods and Cases*. Springer-Verlag, 2008
- [2] B.OKSENDAL: *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 2003
- [3] G.ALSMEYER: *Wahrscheinlichkeitstheorie*. Institut für mathematische Statistik, Fachbereich Mathematik der WWU Münster, 2007
- [4] P.E.KLOEDEN, E.PLATEN, H.SCHURZ: *Numerical Solution of SDE Through Computer Experiments*. Springer-Verlag, 2003
- [5] T.BJÖRK: *Arbitrage Theory in Continuous Time*. Oxford University Press, 2009
- [6] P.GLASSERMAN: *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, 2004
- [7] I.KARATZAS, S.SHREVE *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 2005
- [8] H.-J.RUNCKEL: *Höhere Analysis - Funktionentheorie und gewöhnliche Differentialgleichungen*. Oldenbourg Verlag, 2000
- [9] R.CONT, P.TANKOV: *Financial Modeling with Jump Processes*. Chapman & Hall/CRC, 2004
- [10] M.C. FU, D.B. MADAN, T. WANG: *Pricing Continuous Asian Options: A Comparison of Monte Carlo and Laplace Transform Inversion Methods*. Journal of Computational Finance, Vol.2, No.2, 49-74, 1999
- [11] J.S.DAGPUNAR: *Simulation and Monte Carlo*. John Wiley & Sons, 2007
- [12] DON L.MC LEISH: *Monte Carlo Simulation & Finance*. John Wiley & Sons, 2005
- [13] R.BUFF: *Uncertain Volatility Models - Theory and Application*. Springer, 2002
- [14] M.GÜNTHER, A.JÜNGEL: *Finanzderivate mit Matlab*. Vieweg+Teubner, 2010
- [15] E. DERMAN, I.KANI: *Riding on a Smile*. Risk Vol.7, 1994