

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

Datum

Unterschrift

Inhaltsverzeichnis

1	Einleitung und Motivation	1
2	Mathematische Grundlagen	3
2.1	Verteilungsfunktion und Verallgemeinerte Inverse	3
2.2	Lebesgue-Räume	6
2.3	Konvergenzarten und Konvergenzsätze	7
2.4	Bedingter Erwartungswert	10
2.5	Stochastische Prozesse und Martingale	13
2.6	Brownsche Bewegung und der Satz von Girsanov	15
2.6.1	Eindimensional	15
2.6.2	Mehrdimensional	18
2.7	Finanzmathematische Grundlagen	19
3	Modellrahmen	23
4	Die Risikomaße Value at Risk und Expected Shortfall	25
4.1	Definition der Risikomaße	25
4.2	Kohärenz	27
4.2.1	Value at Risk	27
4.2.2	Expected Shortfall	29
4.3	Weitere Eigenschaften des Value at Risk und Expected Shortfall	32
5	Monte Carlo Simulation	36
5.1	Klassische Monte Carlo Simulation	37
5.2	Simulation von Zufallsvariablen	39
5.3	Simulation von Stichprobenpfade	41
5.3.1	Brownsche Bewegung eindimensional	41
5.3.2	Brownsche Bewegung mit Drift	43
5.3.3	Geometrische Brownsche Bewegung	43
5.3.4	Mehrdimensionale Brownsche Bewegung	44
5.4	Methoden zur Varianzreduktion	45
5.4.1	Antithetische Zufallsvariable	45
5.4.2	Kontrollierende Zufallsvariable	49
5.4.3	Importance Sampling	52
5.4.4	Stratified Sampling	53
5.4.5	Latin Hypercube Sampling	62
6	Monte Carlo Methoden in der Versicherungsmathematik	65
6.1	Nested Monte Carlo Simulationen	65
6.2	Least Square Monte Carlo und replizierendes Portfolio	67
6.2.1	Regressionsmodell für LSMC	67
6.2.2	Regressionsmodell für RP	70
6.2.3	Vor- und Nachteile des RP gegenüber der LSMC	74

7	Replizierendes Portfolio	78
7.1	Dynamisches replizierendes Portfolio	78
7.2	Statisches Replizierendes Portfolio	81
7.3	Analyse des Approximationsfehler	82
8	MC-Methoden angewendet auf RP	85
8.1	Anwendung von Varianzreduzierenden Methoden	89
9	Beispiele	91
9.1	Geometrische Brownsche Bewegung	92
9.2	Europäische Call Option	98
9.3	Asiatische Put Option	101
9.4	Portfolio bestehend aus zwei Finanzderivaten	106
10	Conclusio	110

1 Einleitung und Motivation

Seit der Einführung des Schweizer Solvenztests (SST), welcher im Januar 2011 in Kraft trat, und dem EU weiten Regulierungsrahmen für die Eigenmittelausstattung von Versicherungsunternehmen Solvency II (Januar 2016), wird die Modellierung der Profit- und Verlustverteilung des Asset-Liability Portfolios auf einem Einjahreszeitraum gefordert. Das ist das Gesamtportfolio des Unternehmens, welches sowohl alle Verbindlichkeiten als auch alle Vermögenswerte beinhaltet. Die Eigenmittelkapitalanforderung wird bei Solvency II durch den Value at Risk von 99.5 % festgelegt, und beim SST wird diese mit dem Expected Shortfall von 99 % berechnet.

Im Fall von Solvency II heißt das, dass ein Versicherungsunternehmen mit einer Wahrscheinlichkeit von 99.5 % allen Zahlungsverbindlichkeiten im darauf folgenden Jahr nachkommen kann. Einen genauen Einblick gibt die Richtlinie [Ric09]. Der Value at Risk sagt jedoch nichts über die Schwere eines möglichen Verlustes aus. Der Expected Shortfall, der beim SST verwendet wird, berechnet sich im Gegensatz zum Value at Risk als Durchschnitt der möglichen Verluste, wenn man eine Insolvenz mit einer Wahrscheinlichkeit von 99 % entgehen möchte. Damit ist der Expected Shortfall ein konservativeres Risikomaß als der Value at Risk. In Kapitel 4 werden diese zwei Risikomaß anhand ihrer Eigenschaften genauer miteinander verglichen.

Für die Berechnung dieser beiden Risikomaße wird die Verteilung des Asset-Liability Portfolios nach einem Jahr benötigt. In der Versicherungsbranche ist es möglich, dass man Zahlungsströme bis zu 40 Jahren simulieren muss, was einen enormen rechnerischen Aufwand bedeutet. Dafür bieten sich verschiedene Methoden an, jedoch gehen die Meisten davon entweder auf Kosten der Genauigkeit oder der Schnelligkeit.

In den meisten Fällen ist eine direkte Berechnung nicht möglich, daher benötigt man die Monte Carlo Simulation, um mit dieser den Einjahresgewinn zu schätzen. Die Monte Carlo Simulation ist eine stochastische Simulation und wird im Kapitel 5 genauer erklärt. Jedoch ist die Konvergenzrate der Monte Carlo Simulation $O\left(N^{-\frac{1}{2}}\right)$, was nicht besonders schnell ist. Die Konvergenzgeschwindigkeit hängt aber auch von der Größe der Varianz des Schätzers ab, daher gibt es Methoden, welche die Varianz reduzieren um die Konvergenz zu verbessern. Ein paar dieser Methoden werden im Kapitel 5.4 vorgestellt.

Es gibt mehrere Monte Carlo Ansätze, welche man für die Schätzung der Verteilung des Wertes des Portfolios nach einem Jahr verwenden kann. Die intuitive Vorgehensweise ist Nested Monte Carlo Simulations, jedoch ist der Rechenaufwand dieser Methode enorm. Einen kurzen Einblick in dieses Ver-

fahren bekommt man in Kapitel 6.1.

Zwei weitere Methoden, welche in der Praxis oft verwendet werden, sind Least Square Monte Carlo (LSMC) und der replizierende Portfolio (RP) Ansatz. Beide werden mittels Regression berechnet. Im Kapitel 6.2.3 werden diese beiden Methoden in Bezug auf den Regressionsfehler miteinander verglichen.

Der Hauptteil dieser Arbeit untersucht die Methode des dynamischen pfadabhängigen replizierenden Portfolios. In der Praxis wird ein statische replizierende Portfolio für die Schätzung des Einjahresgewinn verwendet, da der Rechenaufwand um dieses zu berechnen sehr gering ist. Diese Methode hat jedoch einen großen Nachteile:

Die Genauigkeit der Schätzung der Risikomaße ist nicht zu vergleichen mit jener des dynamischen pfadabhängigen Portfolios.

Diese Behauptung wird in den Beispielen, welche am Ende der Arbeit in Kapitel 9 bearbeitet werden, gefestigt. Die Simulation der Beispiele wurde in Matlab vorgenommen. Es werden die geschätzten Risikomaße des statischen replizierenden Portfolios mit jenen verglichen, welche mit dem dynamischen replizierenden Portfolio berechnet werden. Dabei werden auch verschiedene Varianzreduktionsmethoden bei der Schätzung der Verteilung des Einjahresgewinn angewendet und den Schätzern der klassischen Monte Carlo Methode gegenüber gestellt.

2 Mathematische Grundlagen

Im ersten Kapitel werden grundlegende Definitionen sowie Sätze angeführt, welche im Hauptteil der Arbeit verwendet werden.

Im folgenden sei $(\Omega, \mathcal{B}, \mathbb{P})$ der zugrundeliegende Wahrscheinlichkeitsraum, das heißt Ω ist eine nicht leere Menge, \mathcal{B} ist eine σ -Algebra bestehend aus Teilmengen von Ω und \mathbb{P} ist ein Wahrscheinlichkeitsmaß. Alle Zufallsvariable X , welche verwendet werden, sind Funktionen von Ω nach \mathbb{R} .

2.1 Verteilungsfunktion und Verallgemeinerte Inverse

Für die Definition der Riskikomaße benötigt man einige Begriffe aus der Wahrscheinlichkeitstheorie. Die folgenden Definitionen und Eigenschaften findet man detaillierter in [Kre18].

Definition 2.1. Sei X eine reellwertige Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{B}, \mathbb{P})$. Die Verteilungsfunktion $F_X : \mathbb{R} \rightarrow [0, 1]$ von X ist definiert durch

$$F_X(x) = \mathbb{P}(X \leq x).$$

Einige wichtige Eigenschaften, welche die Verteilungsfunktion erfüllt sind in folgendem Satz angegeben, der in [Kre18] bewiesen wird:

Satz 2.1. Die Verteilungsfunktion F_X einer Zufallsvariablen X ist monoton wachsend und rechtsstetig. Weiter gilt

$$\lim_{x \rightarrow -\infty} F_X(x) = 0, \text{ und } \lim_{x \rightarrow \infty} F_X(x) = 1.$$

Da eine Verteilungsfunktion nur monoton steigend ist und nicht streng monoton steigend, ist es im allgemeinen nicht möglich die Inverse der Verteilungsfunktion zu berechnen.

Im Buch [Kre18] werden daher die Verallgemeinerten Inversen $F_{\pm}^{\leftarrow} : \mathbb{R} \rightarrow \bar{\mathbb{R}}$ einer monoton steigenden, rechtsstetige Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ durch

$$F_+^{\leftarrow}(\alpha) = \inf \{x \in \mathbb{R} \mid F(x) > \alpha\} \quad (1)$$

und

$$F_-^{\leftarrow}(\alpha) = \inf \{x \in \mathbb{R} \mid F(x) \geq \alpha\} \quad (2)$$

definiert. Ist F in einer Umgebung von α , für $\alpha \in [0, 1]$, invertierbar mit Umkehrfunktion F^{-1} , dann gilt offensichtlich $F_+^{\leftarrow}(\alpha) = F_-^{\leftarrow}(\alpha) = F^{-1}(\alpha)$. In Kapitel 4 werden die beiden Risikomaße Value at Risk und Expected Shortfall über verallgemeinerte inverse Funktionen definiert. In den folgenden Hilfssätzen werden einige Eigenschaften der verallgemeinerten Inversen angegeben, welche direkt auf die Risikomaße übertragen werden können oder auf anderer Weise benötigt werden.

Lemma 2.1. Für F_{\pm}^{\leftarrow} gelten folgende Aussagen:

1. Sowohl F_{+}^{\leftarrow} als auch F_{-}^{\leftarrow} sind monoton wachsend.
2. Für $\alpha \in [0, 1]$ gilt $F_{-}^{\leftarrow} \leq F_{+}^{\leftarrow}$.
3. Für $0 \leq \alpha < \beta \leq 1$ gilt $F_{+}^{\leftarrow}(\alpha) \leq F_{-}^{\leftarrow}(\beta)$.
4. Für $0 < \alpha < 1$ gilt $-\infty < F_{-}^{\leftarrow}(\alpha) \leq F_{+}^{\leftarrow}(\alpha) < \infty$.

Werden also die Definitionsbereiche von F_{-}^{\leftarrow} und F_{+}^{\leftarrow} auf das offene Intervall $(0, 1)$ eingeschränkt, dann sind beide Einschränkungen reellwertig.

Beweis. 1. Für $0 \leq \alpha < \beta \leq 1$ gilt $\{x \in \mathbb{R} \mid F(x) > \beta\} \subset \{x \in \mathbb{R} \mid F(x) > \alpha\}$ und damit $\inf \{x \in \mathbb{R} \mid F(x) > \beta\} \geq \inf \{x \in \mathbb{R} \mid F(x) > \alpha\}$ also $F_{+}^{\leftarrow}(\alpha) \leq F_{+}^{\leftarrow}(\beta)$.

Für F_{-}^{\leftarrow} folgt die Behauptung analog.

2. Da $\{x \in \mathbb{R} \mid F(x) > \alpha\} \subset \{x \in \mathbb{R} \mid F(x) \geq \alpha\}$ folgt $F_{+}^{\leftarrow}(\alpha) \geq F_{-}^{\leftarrow}(\alpha)$.
3. Für $0 \leq \alpha < \beta \leq 1$ gilt $\{x \in \mathbb{R} \mid F(x) \leq \beta\} \subset \{x \in \mathbb{R} \mid F(x) > \alpha\}$ und damit $F_{-}^{\leftarrow}(\beta) \geq F_{+}^{\leftarrow}(\alpha)$.
4. Sei $0 < \alpha < 1$. Es gilt $F(x) \searrow 0$ für $x \rightarrow -\infty$, daher existiert $x_0 \in \mathbb{R}$ mit $F(x_0) < \alpha$, also $x_0 \notin \{x \in \mathbb{R} \mid F(x) \geq \alpha\}$ und damit $x_0 \leq F_{-}^{\leftarrow}(\alpha)$. Wegen $F(x) \nearrow 1$ für $x \rightarrow \infty$ gilt, dass $x_1 \in \mathbb{R}$ existiert mit $F(x_1) > \alpha$, daher ist $\{x \in \mathbb{R} \mid F(x) > \alpha\}$ nicht leer, und es gilt $x_1 \geq F_{+}^{\leftarrow}(\alpha)$. □

Für eine stetige monoton wachsende Verteilungsfunktion gelten zusätzlich folgende Eigenschaften.

Lemma 2.2. Sei F eine stetige, monoton wachsende Funktion, dann gilt

1. F_{-}^{\leftarrow} ist streng monoton steigend. Daraus folgt $F(F_{-}^{\leftarrow}(y)) = y$.
2. $\mathbb{P}(F_{-}^{\leftarrow}(F(x)) = x) = 1$, also $F_{-}^{\leftarrow}(F(x)) = x$ fast sicher.

Beweis. 1. Sei $\alpha < \beta$, $x = F_{-}^{\leftarrow}(\alpha) = \inf \{z \in \mathbb{R} \mid F(z) \leq \alpha\}$ und $y = F_{-}^{\leftarrow}(\beta) = \inf \{z \in \mathbb{R} \mid F(z) \leq \beta\}$. Da F stetig ist, gilt $F(x) = \alpha$ und $F(y) = \beta$. Aus der Monotonie von F und $\alpha < \beta$ folgt $x < y$ also die Behauptung.

2.

$$\mathbb{P}(F_{-}^{\leftarrow}(F(x)) = x) = \mathbb{P}(F(F_{-}^{\leftarrow}(F(x))) = F(x))$$

aus 1. folgt nun direkt

$$\mathbb{P}(F(x) = F(x)) = 1$$

□

Mit Hilfe des folgenden Satzes ist es möglich eine wichtige Eigenschaft verallgemeinerter inverser Funktionen zu beweisen, welche angewendet werden kann um Stichproben aus beliebigen Verteilungsfunktionen zu simulieren.

Satz 2.2. *Sei $F : \mathbb{R} \rightarrow \mathbb{R}$ eine monoton wachsende, rechtsstetige Funktion mit verallgemeinerter Inversen F_-^{\leftarrow} . Dann gilt für jedes $\alpha \in (0, 1)$*

$$\alpha \leq F(x) \iff F_-^{\leftarrow}(\alpha) = x. \quad (3)$$

Beweis. „ \Rightarrow “ Sei $F(x) \geq \alpha$, dann ist $x \in \{z \in \mathbb{R} \mid F(z) \leq \alpha\}$, also ist $x \leq \inf \{z \in \mathbb{R} \mid F(z) \leq \alpha\} = F_-^{\leftarrow}(\alpha)$.

„ \Leftarrow “ Sei $F(x) < \alpha$ dann gilt $x \notin \{z \in \mathbb{R} \mid F(z) \leq \alpha\}$, da F rechtsstetig ist, folgt $x < \inf \{z \in \mathbb{R} \mid F(z) \leq \alpha\} = F_-^{\leftarrow}(\alpha)$. Was nun bedeutet $F_-^{\leftarrow}(\alpha) \leq x$ und daher gilt auch $F(x) \geq \alpha$. \square

Damit kann nun dieser Satz leicht gefolgert werden, welcher es ermöglicht aus standardgleichverteilten Zufallszahlen eine Stichprobe jeder beliebigen Verteilung zu berechnen, deren verallgemeinerte Inverse bekannt ist.

Satz 2.3. *Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion mit verallgemeinerter Inversen Funktion F_-^{\leftarrow} . Weiters sei U eine gleichverteilte Zufallsvariable auf $(0, 1)$, das heißt $\mathbb{P}(U \leq u) = u$, dann ist F die Verteilungsfunktion von $F_-^{\leftarrow}(U)$.*

Beweis. Mit (3) gilt

$$\mathbb{P}(F_-^{\leftarrow}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x)$$

\square

Die verallgemeinerten Inversen von Verteilungsfunktionen werden als Quantilfunktionen bezeichnet.

Definition 2.2. *Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{B}, \mathbb{P})$ mit Verteilungsfunktion F . Für $\alpha \in (0, 1)$ werden*

$$q_X^-(\alpha) = F_-^{\leftarrow}(\alpha)$$

das untere Quantil von X und

$$q_X^+(\alpha) = F_+^{\leftarrow}(\alpha)$$

das obere Quantil von X genannt.

Folgende wichtige Eigenschaften der Quantilfunktion werden später noch öfter benötigt.

Satz 2.4. *Sei $-\infty \leq a < b \leq \infty$ und $f : \mathbb{R} \rightarrow (a, b)$ streng monoton wachsend und bijektiv, dann gilt*

$$q_{f(X)}^-(\alpha) = f(q_X^-(\alpha)), \quad q_{f(X)}^+(\alpha) = f(q_X^+(\alpha))$$

Beweis. Sei $a < x < b$, dann gilt

$$\begin{aligned} \{x \in \mathbb{R} \mid \mathbb{P}(f(X) \leq x) \geq \alpha\} &= \{x \in \mathbb{R} \mid \mathbb{P}(X \leq f^{-1}(x)) \geq \alpha\} \\ &= \{f(y) \in \mathbb{R} \mid \mathbb{P}(X \leq y) \geq \alpha\} \end{aligned}$$

woraus

$$\begin{aligned} q_{f(X)}^-(\alpha) &= \inf \{f(y) \in \mathbb{R} \mid \mathbb{P}(X \leq y) \geq \alpha\} \\ &= f(\inf \{y \in \mathbb{R} \mid \mathbb{P}(X \leq y) \geq \alpha\}) = f(q_X^-(\alpha)) \end{aligned}$$

folgt, denn als streng monoton wachsende, bijektive Funktion ist f stetig.

Zweite Gleichheit folgt analog. \square

Aus diesem Satz folgt direkt:

Korollar 2.1. *Sei X eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Für $\mu \in \mathbb{R}$ und $\sigma > 0$ gilt*

$$q_{\mu+\sigma X}^-(\alpha) = \mu + \sigma q_X^-(\alpha), \quad q_{\mu+\sigma X}^+(\alpha) = \mu + \sigma q_X^+(\alpha)$$

2.2 Lebesgue-Räume

Es werden nun die Lebesgue-Räume eingeführt, auf welchen man einen Norm definieren kann. Die Definitionen stammen aus [Kle13]. Für $f : \Omega \rightarrow \bar{\mathbb{R}}$ definiert man

$$\|f\|_{L^p(\mathbb{P})} = E_{\mathbb{P}} [|f|^p]^{\frac{1}{p}} = \left(\int_{\mathbb{R}} |f|^p d\mathbb{P} \right)^{\frac{1}{p}}$$

für $p \in (0, \infty)$. Weiter sei

$$\mathcal{L}^p(\Omega, \mathcal{B}, \mathbb{P}) = \mathcal{L}^p(\mathbb{P}) = \left\{ f : \Omega \rightarrow \bar{\mathbb{R}} \mid f \text{ ist messbar und } \|f\|_{L^p(\mathbb{P})} < \infty \right\}$$

und

$$\mathcal{N} = \{f \mid f \text{ ist messbar und } f = 0 \text{ } \mathbb{P}\text{-fast überall}\}.$$

Da auf dem Raum $\mathcal{L}^p(\mathbb{P})$ $\|\cdot\|_{L^p(\mathbb{P})}$ keine Norm ist da $\|f\|_{L^p(\mathbb{P})} = 0$ für alle f für welche $f = 0$ \mathbb{P} -fast sicher, definiert man sich den Raum L^p , auf welchem dies eine Norm ist, wie folgt:

Definition 2.3. *Für jedes $p \in [1, \infty)$ definiere*

$$L^p(\mathbb{P}) = \mathcal{L}^p(\mathbb{P}) \setminus \mathcal{N} = \{\bar{f} := f + \mathcal{N} \mid f \in \mathcal{L}^p\}.$$

Für $\bar{f} \in L^p(\mathbb{P})$ setzt man $\|\bar{f}\|_{L^p(\mathbb{P})} = \|f\|_{L^p(\mathbb{P})}$ für ein $f \in \bar{f}$.

Die Funktionen werden in $L^p(\mathbb{P})$ in Äquivalenzklassen zusammengefasst, damit ist $\|\cdot\|_{L^p(\mathbb{P})}$ eine Norm, und es gilt, dass $\|\bar{f}\|_{L^p(\mathbb{P})}$ nicht von der Wahl des Repräsentanten abhängt.

Des weiteren gilt folgender Satz, welcher in [Kle13] bewiesen wird:

Satz 2.5. Für jedes $p \in (0, \infty)$ ist $(L^p(\mathbb{P}), \|\cdot\|_{L^p(\mathbb{P})})$ ein Banachraum.

Insbesondere ist $(L^p(\mathbb{P}), \|\cdot\|_{L^p(\mathbb{P})})$ für $p = 2$ ein Hilbertraum mit Skalarprodukt

$$\langle X, Y \rangle = E_{\mathbb{P}}[XY]$$

wobei X und Y Zufallsvariablen sind.

Für das Skalarprodukt in $L^2(\mathbb{P})$ gilt die Cauchy-Schwarzsche Ungleichung:

$$\langle X, Y \rangle_{L^2(\mathbb{P})}^2 \leq \|X\|_{L^2(\mathbb{P})}^2 \|Y\|_{L^2(\mathbb{P})}^2, \quad (4)$$

denn für $\|X\|_{L^2(\mathbb{P})}^2 \neq 0$ (für $\|X\|_{L^2(\mathbb{P})}^2 = 0$ ist die Aussage trivial) setze $a = \frac{\langle X, Y \rangle_{L^2(\mathbb{P})}}{\|X\|_{L^2(\mathbb{P})}^2} = \frac{E_{\mathbb{P}}[XY]}{E_{\mathbb{P}}[X^2]}$, dann gilt

$$\begin{aligned} 0 &\leq \|Y - aX\|_{L^2(\mathbb{P})}^2 \|X\|_{L^2(\mathbb{P})}^2 = E_{\mathbb{P}}[(Y - aX)^2] E_{\mathbb{P}}[X^2] \\ &= E_{\mathbb{P}}[Y^2] E_{\mathbb{P}}[X^2] - E_{\mathbb{P}}[XY]^2 = \|Y\|_{L^2(\mathbb{P})}^2 \|X\|_{L^2(\mathbb{P})}^2 - \langle X, Y \rangle_{L^2(\mathbb{P})}^2. \end{aligned}$$

2.3 Konvergenzarten und Konvergenzsätze

Die folgenden Definitionen und Sätze sind die Grundlage der Monte Carlo Simulation. Der Großteil der Beweise und genauere Erläuterungen der hier nur kurz angeführten Begriffe und Sätze können in [BnHH⁺16] Kapitel Konvergenzbegriffe und Grenzwertsätze - Stochastik für große Stichproben nachgelesen werden.

Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{B}, \mathbb{P})$ mit Verteilungsfunktionen F_n , und sei X eine Zufallsvariable mit Verteilungsfunktion F . Es werden nun die Definitionen einiger Konvergenzbegriffe angegeben:

- Fast sichere Konvergenz:
Man sagt die Folge $(X_n)_{n \in \mathbb{N}}$ konvergiert fast sicher gegen X , wenn

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

was bedeutet

$$\mathbb{P}\left(\left\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

- Konvergenz in Wahrscheinlichkeit bzw. stochastische Konvergenz:
Es gilt $(X_n)_{n \in \mathbb{N}}$ konvergiert in Wahrscheinlichkeit gegen die Zufallsvariable X , wenn

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$$

- Konvergenz in der p -Norm:

Für $0 < p < \infty$ gilt $(X_n)_{n \in \mathbb{N}}$ konvergiert in der p -Norm gegen X , wenn sowohl X_n für alle $n \in \mathbb{N}$ als auch X endlichen p -ten Moment haben und

$$\lim_{n \rightarrow \infty} E [|X_n - X|^p] = 0.$$

- Konvergenz in Verteilung:

Eine Folge von Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$ konvergiert in Verteilung gegen eine Zufallsvariable X , wenn $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ für alle Stetigkeitsstellen $x \in \mathbb{R}$. Was, wie in [Gla04] angeführt, äquivalent ist zu

$$E[f(X_n)] \rightarrow E[f(X)]$$

für alle beschränkten stetigen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ und zur punktweisen Konvergenz der charakteristischen Funktion.

Die Konvergenz in Verteilung wird auch schwache Konvergenz genannt.

Für Zufallsvektoren definiert man die Konvergenzbegriffe fast sichere Konvergenz, Konvergenz in Wahrscheinlichkeit und Konvergenz in der p -Norm, wie es in [Gla04] auch kurz erwähnt wird, über die Komponenten des Vektors. Was zum Beispiel für die fast sichere Konvergenz bedeutet, dass eine Folge von Zufallsvektoren $(X_n)_{n \in \mathbb{N}}$, wobei $X_n = (X_n^1, \dots, X_n^d) \in \mathbb{R}^d$, konvergiert genau dann fast sicher gegen einen Zufallsvektor $X \in \mathbb{R}^d$, wenn die einzelnen Komponenten fast sicher gegen die Komponenten von X konvergieren.

Die Konvergenz in Verteilung von Vektoren wird durch die Konvergenz der Erwartung aller beschränkten stetigen Funktionen definiert, das heißt eine Folge d -dimensionaler Zufallsvektoren X_1, X_2, \dots konvergiert genau dann in Verteilung, wenn

$$\lim_{n \rightarrow \infty} E[f(X_n)] = E[f(X)]$$

für alle beschränkten stetigen Funktionen f .

Der folgende Satz gibt etwas Einblick wie die unterschiedlichen Konvergenzbegriffe miteinander in Verbindung gebracht werden können. Man kann sehen, dass die Konvergenz in Verteilung der „schwächste“ Konvergenzbegriff ist, was bedeutet, dass diese Konvergenz von allen anderen impliziert wird.

Über einen Zusammenhang zwischen Konvergenz in der p -Norm und fast sichere Konvergenz gibt es keine Aussage.

Dieser Satz ist eine Zusammenfassung von mehreren Sätzen, welche in [BnHH⁺16] bewiesen werden.

- Satz 2.6.**
1. *Fast sichere Konvergenz \Rightarrow Konvergenz in Wahrscheinlichkeit*
 2. *Konvergenz in Wahrscheinlichkeit \Rightarrow Es existiert eine Teilfolge, sodass die Teilfolge fast sicher konvergiert.*
 3. *Konvergenz in der p -Norm \Rightarrow Konvergenz in Wahrscheinlichkeit*
 4. *Konvergenz in Wahrscheinlichkeit \Rightarrow Konvergenz in Verteilung*

Die nächsten Sätze sind wichtige Sätze, auf denen die Theorie der Monte Carlo Simulation aufbaut. Der folgende Satz, welcher das starke Gesetz der großen Zahlen folgert, wobei es dabei darum geht, unter welchen Voraussetzungen das arithmetische Mittel einer Folge von Zufallsvariablen gegen den Erwartungswert der Zufallsvariablen konvergiert. Hier wird das starke Gesetz der großen Zahlen nach Kolmogorov angeführt. Es gibt jedoch einige weitere Versionen dieses Satzes.

Satz 2.7. *Es sei $(X_n)_{n \geq 1}$ eine Folge unabhängiger identisch verteilter Zufallsvariablen. Folgende Aussagen sind äquivalent*

1. $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow X$ fast sicher für eine Zufallsvariable X .
2. $E[X_1] < \infty$

In diesem Fall gilt $X = E[X_1]$ fast sicher, daher genügt $(X_n)_{n \geq 1}$ dem starken Gesetz der großen Zahlen, das heißt $\bar{X}_n \rightarrow E[X_1]$ fast sicher.

Es gibt auch das schwache Gesetz der großen Zahlen, welches eine analoge Aussage für die Konvergenz in der Wahrscheinlichkeit beinhaltet.

Satz 2.8. *Es sei $(X_n)_{n \geq 1}$ eine Folge paarweiser unabhängiger identisch verteilter Zufallsvariablen, welche endlichen zweiten Moment haben, also $E[X^2] < \infty$. Dann erfüllt $(X_n)_{n \geq 1}$ das schwache Gesetz der großen Zahlen, was bedeutet \bar{X}_n konvergiert in Wahrscheinlichkeit gegen $E[X]$.*

Das folgende Lemma beinhaltet einige Eigenschaften der Konvergenz in Verteilung und wird als Lemma von Slutsky bezeichnet.

Lemma 2.3. *Es seien X, X_1, X_2, \dots und Y_1, Y_2, \dots Zufallsvariablen wobei $X_n \rightarrow X$ in Verteilung und $Y_n \rightarrow a$ in Wahrscheinlichkeit für ein $a \in \mathbb{R}$. Dann gelten folgende Aussagen:*

1. $X_n + Y_n \rightarrow X + a$ in Verteilung.
2. $X_n Y_n \rightarrow aX$ in Verteilung.

Der zentrale Grenzwertsatz besagt, dass unter bestimmten Voraussetzungen das standardisierte arithmetische Mittel für große n asymptotisch standardnormalverteilt ist. Der Satz wird hier eindimensional und mehrdimensional angegeben.

Satz 2.9. Sei $(X_n)_{n \geq 1}$ eine Folge unabhängiger und identisch verteilter reeller Zufallsvariablen mit $E[X_n] = \mu$, $\text{Var}[X_n] = \sigma^2$ und $0 < \sigma^2 < \infty$. Bezeichnet $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ das arithmetische Mittel der X_i , dann gilt

$$\lim_{n \rightarrow \infty} \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1).$$

Satz 2.10. Sei $(X_n)_{n \geq 1}$ eine Folge unabhängiger, identisch verteilter d -dimensionaler Zufallsvektoren mit $E[X_n] = \mu \in \mathbb{R}^d$ und $\text{Var}[X_n] = \Sigma \in \mathbb{R}^{d \times d}$, dann gilt

$$\lim_{n \rightarrow \infty} \sqrt{n} (\bar{X}_n - \mu) \sim N_d(0, \Sigma).$$

2.4 Bedingter Erwartungswert

Ein weiterer wichtiger Begriff, welcher eine große Rolle in der Finanzmathematik spielt ist die bedingte Erwartung.

In diesem Kapitel werden die Definitionen und Eigenschaften aus dem Buch [Kle13] verwendet.

Der Begriff der bedingten Erwartung, definiert die Änderung der Erwartung oder auch Wahrscheinlichkeit, wenn bereits ein paar Informationen über den Ausgang eines Zufallsexperimentes bekannt sind.

Beispiel 2.1. Nimmt man einen fairen Würfel und betrachte die zwei Zufallsereignisse:

$A =$ eine ungerade Zahl wird gewürfelt, also $A = \{1, 3, 5\}$

$B =$ Gewürfelte Zahl ist größer oder gleich 4, also $B = \{4, 5, 6\}$

Dann gilt $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$.

Weiß man nun bereits bevor man würfelt, dass B eintritt, dann ändert sich die Wahrscheinlichkeit von A . Es gilt $\mathbb{P}\{A|B\}$ ist die Wahrscheinlichkeit, dass eine 5 gewürfelt wird, wenn es möglich ist 4, 5 oder 6 zu würfeln, also $\mathbb{P}\{A|B\} = \frac{1}{3}$.

Die einfache Berechnung dieser bedingten Wahrscheinlichkeit wird in [Kle13] durch folgende Definition der bedingten Erwartung und bedingten Wahrscheinlichkeit verallgemeinert:

Definition 2.4. Eine Zufallsvariable Y heißt bedingte Erwartung von X gegeben \mathcal{F} , $Y = E[X|\mathcal{F}]$, wobei \mathcal{F} eine Unter- σ -Algebra von \mathcal{B} ist, falls gilt:

1. Y ist \mathcal{F} -messbar.
2. Für jedes $A \in \mathcal{F}$ gilt $E[X \mathbf{1}_A] = E[Y \mathbf{1}_A]$,

wobei $\mathbf{1}$ die Indikatorfunktion bezeichnet. Für $B \in \mathcal{B}$ heißt $\mathbb{P}(B|\mathcal{F}) := E[\mathbf{1}_B|\mathcal{F}]$ die bedingte Wahrscheinlichkeit von B gegeben \mathcal{F} .

Der folgende Satz garantiert die Existenz der bedingten Erwartung.

Satz 2.11. $E[X|\mathcal{F}]$ existiert und ist eindeutig (bis auf Gleichheit fast sicher).

Der Satz folgt aus dem Satz von Radon-Nikodym und wird wie auch der Beweis des folgenden Satzes in [Kle13] genau ausgeführt.

Der folgende Satz ist eine Auflistung der Eigenschaften der bedingten Erwartung, dabei ist zu beachten, dass alle Gleichheiten nur als fast sichere Gleichheiten zu verstehen sind.

Satz 2.12. Seien X und Y Zufallsvariablen aus dem $\mathcal{L}^1(\mathbb{P})$, und seien $\mathcal{G} \subset \mathcal{F} \subset \mathcal{B}$ σ -Algebren. Dann gilt

1. *Linearität:*

$$E[\lambda X + Y|\mathcal{F}] = \lambda E[X|\mathcal{F}] + E[Y|\mathcal{F}]$$

für $\lambda \in \mathbb{R}$

2. *Monotonie:*

Ist $X \leq Y$ fast sicher, so ist auch $E[X|\mathcal{F}] \leq E[Y|\mathcal{F}]$ fast sicher.

3. Ist $E[|XY|] < \infty$ und ist Y \mathcal{F} -messbar, dann ist

$$E[XY|\mathcal{F}] = Y E[X|\mathcal{F}] \quad \text{und} \quad E[Y|\mathcal{F}] = E[Y|Y] = Y$$

4. *Glättungseigenschaft oder Turmeigenschaft:*

$$E[E[X|\mathcal{F}]|\mathcal{G}] = E[E[X|\mathcal{G}]|\mathcal{F}] = E[X|\mathcal{G}]$$

5. *Dreiecksungleichung:*

$$E[|X||\mathcal{F}] \geq |E[X|\mathcal{F}]|$$

6. *Unabhängigkeit:*

Sind $\sigma(X)$ und \mathcal{F} unabhängig, dann ist $E[X|\mathcal{F}] = E[X]$.

7. *Majorisierte Konvergenz:*

Ist $Y \geq 0$ und $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen mit $|X_n| \leq Y$ für alle $n \in \mathbb{N}$, sowie $X_n \rightarrow X$ fast sicher für $n \rightarrow \infty$, dann gilt

$$\lim_{n \rightarrow \infty} E[X_n|\mathcal{F}] = E[X|\mathcal{F}] \quad \text{fast sicher und in } \mathcal{L}^1(\mathbb{P})$$

Bemerkung 2.1. • Man kann die σ -Algebra als die Vorinformation betrachten, die man über das Zufallsexperiment hat. Daher kann der bedingte Erwartungswert als beste Vorhersage für den Wert von X betrachtet werden, wenn die Information aus der σ -Algebra \mathcal{F} bekannt ist.

- Sind also X und \mathcal{F} unabhängig, erhält man keine weiteren Informationen über X und daher ist die beste Vorhersage, die man über den Wert von X treffen kann, der Erwartungswert selbst. Das ist genau die Aussage des Punktes Unabhängigkeit im Satz zuvor.
- Ist Y \mathcal{F} -messbar, dann ist der Wert von Y bereits bekannt, und die Vorhersage ist dann Y selbst. Was der dritte Punkt des Satzes beinhaltet.

Eine weitere Eigenschaft, welche bei der Portfolio Replikation verwendet wird, ist die Betrachtung des bedingten Erwartungswertes als orthogonale Projektion, welche in folgenden Korollar bewiesen wird.

Korollar 2.2. Sei $\mathcal{F} \subset \mathcal{B}$ eine σ -Algebra und X eine Zufallsvariable mit $E[X^2] < \infty$. Dann ist $E[X|\mathcal{F}]$ die orthogonale Projektion von X auf $\mathcal{L}^2(\mathbb{P})$. Es gilt also für jedes \mathcal{F} -messbare Y mit $E[Y^2] < \infty$

$$E[(X - Y)^2] \geq E[(X - E[X|\mathcal{F}])^2]$$

mit Gleichheit genau dann, wenn $Y = E[X|\mathcal{F}]$.

Beweis. Es wird zuerst angenommen $E[E[X|\mathcal{F}]^2] < \infty$. Direkt im Anschluss wird gezeigt, dass das immer der Fall ist, wenn $E[X^2] < \infty$. Sei Y eine beliebige \mathcal{F} -messbare, quadratisch integrierbare Zufallsvariable. Dann gilt mit der Unabhängigkeitseigenschaft und der Glättungseigenschaft

$$E[XY] = E[E[XY|\mathcal{F}]] = E[YE[X|\mathcal{F}]]$$

und für $Y = E[X|\mathcal{F}]$ gilt weiter

$$E[XE[X|\mathcal{F}]] = E[E[X|\mathcal{F}]^2].$$

Daraus folgt

$$\begin{aligned} & E[(X - Y)^2] - E[(X - E[X|\mathcal{F}])^2] \\ &= -2E[XY] + E[Y^2] + 2E[XE[X|\mathcal{F}]] - E[E[X|\mathcal{F}]^2] \\ &= E[(E[X|\mathcal{F}] - Y)^2] \geq 0. \end{aligned}$$

Gleichheit gilt genau dann, wenn $E[(E[X|\mathcal{F}] - Y)^2] = 0$, also wenn $Y = E[X|\mathcal{F}]$ fast sicher.

Nun wird noch bewiesen, dass $E[E[X|\mathcal{F}]^2] < \infty$, wenn $E[X^2] < \infty$. Dazu betrachte man die Zufallsvariable $|X| \wedge N = \min\{|X|, N\}$ für $N \in \mathbb{N}$. Offensichtlich gilt $E[|X| \wedge N] \leq N$. Mit dem bereits gezeigten

und mit Hilfe der Ungleichung $a^2 \leq 2(a - b)^2 + b^2$ für $a, b \in \mathbb{R}$, welche aus $0 \leq (a - 2b)^2$ folgt, gilt

$$\begin{aligned} E \left[E[|X| \wedge N | \mathcal{F}]^2 \right] &\leq 2E \left[((|X| \wedge N) - E[|X| \wedge N | \mathcal{F}])^2 \right] + 2E \left[(|X| \wedge N)^2 \right] \\ &= 4E \left[(|X| \wedge N)^2 \right] - 2E \left[E[|X| \wedge N | \mathcal{F}]^2 \right] \\ &\leq 4E \left[(|X| \wedge N)^2 \right] \leq 4E[X^2]. \end{aligned}$$

Damit erhält man mit der Majorisierten Konvergenz und der Monotonie

$$E[|X| \wedge N | \mathcal{F}] \nearrow E[|X| | \mathcal{F}].$$

Des weiteren erhält man mit der Dreieckungleichung und den Satz über die monotone Konvergenz für Integrale die Behauptung:

$$\begin{aligned} E \left[E[X | \mathcal{F}]^2 \right] &\leq E \left[E[|X| | \mathcal{F}]^2 \right] \\ &= \lim_{N \rightarrow \infty} E \left[E[|X| \wedge N | \mathcal{F}]^2 \right] \leq 4E[X^2] \end{aligned}$$

□

2.5 Stochastische Prozesse und Martingale

In der Finanzmathematik werden Aktienkurse etc. über stochastische Prozesse angegeben. Dafür werden nun einige grundlegende Definitionen, welche stochastische Prozesse betreffen hier formuliert. Diese Definitionen sind aus dem Buch [Kle13].

Zuerst eine ganz allgemeine Definition für stochastische Prozesse:

Definition 2.5. Sei $I \subset \mathbb{R}$. Eine Familie von Zufallsvariablen $X = (X_t)_{t \in I}$ auf $(\Omega, \mathcal{B}, \mathbb{P})$ mit Werten in (E, \mathcal{E}) heißt stochastischer Prozess mit Zeitbereich I und Zustandsraum E . Hier ist (E, \mathcal{E}) ein Messraum.

Definition 2.6. Ein stochastischer Prozess $X = (X_t)_{t \in I}$ mit Werten in E heißt

1. reellwertig, falls $E = \mathbb{R}$.
2. Prozess mit unabhängigen Zuwächsen, falls X reellwertig ist und für jedes $n \in \mathbb{N}$ und alle $t_0, \dots, t_n \in I$ mit $t_0 < t_1 < \dots < t_n$ gilt

$$(X_{t_i} - X_{t_{i-1}})_{i=1, \dots, n}$$

sind unabhängig.

3. integrierbar (bzw. quadratisch integrierbar), falls X reellwertig ist und $E[|X_t|] < \infty$ (bzw. $E[X_t^2] < \infty$) für jedes $t \in I$ gilt.

Ein weiterer wichtiger Begriff der oft in der Finanzmathematik vorkommt ist die Filtration, welche wie folgt definiert wird:

Definition 2.7. Eine Familie $\mathcal{F} = (\mathcal{F}_t)_{t \in I}$ von σ -Algebren mit $\mathcal{F}_t \subset \mathcal{B}$ für jedes $t \in I$ heißt Filtration, falls $\mathcal{F}_s \subset \mathcal{F}_t$ für alle $s, t \in I$ für die $s \leq t$ gilt.

Definition 2.8. Ein stochastischer Prozess $X = (X_t)_{t \in I}$ heißt adaptiert an die Filtration \mathcal{F} , falls X_t bezüglich \mathcal{F}_t messbar ist für jedes $t \in I$. Gilt $\mathcal{F}_t = \sigma(X_s, s \leq t)$ für jedes $t \in I$, dann nennt man \mathcal{F} die von X erzeugte Filtration.

Definition 2.9. Ein stochastischer Prozess $X = (X_n)_{n \in \mathbb{N}_0}$ heißt vorher-sagbar bezüglich der Filtration $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{N}_0}$, falls X_0 konstant ist und für jedes $n \in \mathbb{N}$ gilt:

X_n ist \mathcal{F}_{n-1} -messbar.

Diese Definition kann für Prozesse und Filtrationen, welche auf beliebige diskrete Teilmengen $I \subset \mathbb{R}$ definiert sind, analog angegeben werden.

Definition 2.10. Eine Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ heißt rechtsstetig falls $\mathcal{F} = \mathcal{F}^+$, wobei $\mathcal{F}_t^+ = \bigcap_{s > t} \mathcal{F}_s$.

Eine Filtration \mathcal{F} heißt \mathbb{P} -vollständig, wenn \mathcal{F}_0 alle \mathbb{P} -Nullmengen enthält, das heißt $\mathcal{N} := \{A \in \mathcal{B} \mid \mathbb{P}(A) = 0\} \subset \mathcal{F}_0$.

Man sagt eine Filtration \mathcal{F} erfüllt die üblichen Bedingungen, wenn \mathcal{F} rechtsstetig und \mathbb{P} -vollständig ist.

Ein weiterer Begriff, welcher eine wesentliche Rolle in der Wahrscheinlichkeitstheorie und auch in der Finanzmathematik hat, ist das Martingal. Es formalisiert die Idee des fairen Spieles, wie es in [Kle13] beschrieben wird. Das bedeutet, dass die Auszahlung im Mittel fair ist. Folgendes Beispiel beschreibt die Idee hinter der Definition des Martingals.

Beispiel 2.2. Betrachtet man ein Glücksspiel bei dem beliebig viele Runden gespielt werden, wobei jede Runde unabhängig von den Ergebnissen der Vor-runden ist. Es wird nun angenommen, dass das Startkapital eine Konstante K ist, dann ist das neue Kapital X_1 nach der ersten Runde

$$X_1 = K + Y_1$$

wobei Y_1 sowohl positiv, negativ als auch Null sein kann. Hat man nun n Runden gespielt, beträgt das neue Kapital

$$X_n = K + \sum_{i=1}^n Y_i$$

Es sind nun n Runden gespielt, das heißt Y_i für $i = 1, \dots, n$ sind bekannt. Spielt man nun die nächste Runde, sollte bei einem fairen Spiel $E[Y_{n+1}] = 0$

gelten. Daraus kann man unter Verwendung der Eigenschaften des bedingten Erwartungswertes folgern, dass

$$E[X_{n+1}|Y_1, \dots, Y_n] = E\left[K + \sum_{i=1}^n Y_i + Y_{n+1}|Y_1, \dots, Y_n\right] = X_n + E[Y_{n+1}] = X_n$$

ist.

Formal definiert man ein Martingal durch:

Definition 2.11. Ein reellwertiger Prozess $X = (X_t)_{t \in I}$ heißt Martingal bezüglich der Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in I}$, wenn er an die Filtration adaptiert ist, $E[|X_t|] < \infty$ für jedes $t \in I$ und

$$E[X_t | \mathcal{F}_s] = X_s$$

für alle $s, t \in I$ mit $s < t$.

2.6 Brownsche Bewegung und der Satz von Girsanov

Zwei Beispiele für Martingale, welche als Grundlage für einfache stochastische Modelle in der Finanzmathematik herangezogen werden, sind die Brownsche Bewegung und ihr Exponentialmartingal. Für die Definition der Brownschen Bewegung benötigt man zuerst die Begriff eines stetigen Pfades, welcher in [Kle13] wie folgt definiert ist:

Definition 2.12. Sei $I \subset \mathbb{R}$ und $X = (X_t)_{t \in I}$ ein stochastischer Prozess mit Werten in einem metrischen Raum (E, d) . Für jedes $\omega \in \Omega$ nennt man die Abbildung von I nach E mit $t \mapsto X_t(\omega)$ einen Pfad von X . Es heißt X hat fast sicher stetige Pfade oder kurz X fast sicher stetig, falls für jedes $\omega \in \Omega$ der Pfad $t \mapsto X_t(\omega)$ stetig ist.

Die folgenden Definitionen sind in beiden Fällen aus dem Buch [Gla04]:

2.6.1 Eindimensional

Definition 2.13. Eine Brownsche Bewegung auf $[0, T]$ ist ein stochastischer Prozess $W = (W_t)_{0 \leq t \leq T}$, welcher folgende Eigenschaften erfüllt:

1. $W_0 = 0$.
2. Die Abbildung $t \mapsto W_t$ ist mit Wahrscheinlichkeit 1 eine stetige Funktion auf $[0, T]$.
3. Die Zuwächse $\{W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_k} - W_{t_{k-1}}\}$ sind unabhängig für alle k und für alle $0 \leq t_0 < t_1 < \dots < t_k \leq T$.
4. $W_t - W_s \sim N(0, t - s)$ für alle $0 \leq s < t \leq T$.

Eine Konsequenz aus 1. und 4. ist

$$W_t \sim N(0, t), \text{ für } 0 \leq t \leq T.$$

Bemerkung 2.2. Die Brownsche Bewegung ist ein Martingal bezüglich der von ihr erzeugten Filtration.

Beispiel 2.3. Ein weiteres Beispiel für ein Martingal ist das Exponentialmartingal der Brownschen Bewegung. Es gilt für ein $\sigma \in \mathbb{R}$ ist $Z = (Z_t)_{t \geq 0}$ ein Martingal, wobei

$$Z_t = \exp\left(\sigma W_t - \frac{\sigma^2}{2}t\right).$$

Ein stochastischer Prozess, welcher über eine Brownsche Bewegung definiert wird, ist die Brownsche Bewegung mit Drift.

Definition 2.14. Für Konstanten $\mu \in \mathbb{R}$ und $\sigma > 0$ nennt man einen Prozess $X = (X_t)_{0 \leq t \leq T}$ eine Brownsche Bewegung mit Drift, wenn

$$\frac{X_t - \mu t}{\sigma} \tag{5}$$

eine Brownsche Bewegung ist. μ wird als Drift und σ^2 als Diffusionskoeffizient bezeichnet.

Man kann X aus einer Brownschen Bewegung W konstruieren, denn es gilt

$$X_t = \mu t + \sigma W_t.$$

Daraus kann gefolgert werden, dass $X_t \sim N(\mu t, \sigma^2 t)$. Des Weiteren ist X die Lösung folgender stochastischer Differentialgleichung

$$dX_t = \mu dt + \sigma dW_t.$$

Der Satz von Girsanov besagt nun, dass eine Brownsche Bewegung mit Drift eine Brownsche Bewegung ist, wenn man zu einem anderen Wahrscheinlichkeitsmaß wechselt. Hier wird nur ein linearer Driftterm betrachtet, eine allgemeinere Formulierung des Satzes findet man in [Gla04].

Satz 2.13. (Satz von Girsanov für Brownsche Bewegungen mit linearem Drift)

Sei W eine Brownsche Bewegung bezüglich \mathcal{F} und $\gamma \in \mathbb{R}$. Durch

$$\frac{d\mathbb{Q}_t}{d\mathbb{P}} = Z_t := \exp\left(\gamma W_t - \frac{\gamma^2}{2}t\right)$$

wird ein zu \mathbb{P} äquivalentes Wahrscheinlichkeitsmaß festgelegt. Die durch

$$X_s = W_s - \gamma s$$

für $s \leq t$ definierte Brownsche Bewegung mit Drift ist bezüglich \mathbb{Q}_t eine Brownsche Bewegung.

Beweis. Z_t ist das Exponentialmartingal der Brownschen Bewegung, daher gilt $E[Z_t] = E[Z_0] = 1$. Damit ist \mathbb{Q}_t ein Wahrscheinlichkeitsmaß, welches äquivalent zu \mathbb{P} ist, da $Z_t > 0$ gilt. Man betrachte nun die Zuwächse $X_u - X_v$ für $0 \leq v \leq u \leq t$ gegeben \mathcal{F}_v . Mit der Formel von Bayes folgt:

$$\begin{aligned} E_{\mathbb{Q}_t} [\exp(ix(X_u - X_v)) | \mathcal{F}_v] &= E_{\mathbb{P}} \left[\exp(ix(X_u - X_v)) \frac{Z_t}{Z_v} | \mathcal{F}_v \right] \\ &= E_{\mathbb{P}} \left[\exp(ix(X_u - X_v)) E \left[\frac{Z_t}{Z_v} | \mathcal{F}_u \right] | \mathcal{F}_v \right] = (\times) \end{aligned}$$

Da die Zuwächse der Brownschen Bewegung unabhängig von \mathcal{F}_v sind und Z ein Martingal ist, erhält man durch einsetzen und Anwendung der charakteristischen Funktion der Normalverteilung:

$$\begin{aligned} (\times) &= E_{\mathbb{P}} \left[\exp \left(ix((B_u - B_v) - \gamma(u - v)) + \gamma(B_u - B_v) - \frac{\gamma^2}{2}(u - v) \right) \right] \\ &= \exp(-ix\gamma(u - v)) - \frac{\gamma^2}{2}(u - v) E_{\mathbb{P}} [\exp((ix + \gamma)(W_u - W_v))] \\ &\quad \stackrel{W_u - W_v \sim N(0, u-v)}{=} \exp \left(\frac{x^2}{2}(u - v) \right) \end{aligned}$$

Was genau die charakteristische Funktion einer Normalverteilung mit Parameter 0 und $u - v$ ist. Es gilt also, dass die Zuwächse unabhängig sind von \mathcal{F}_v und normalverteilt. Daher ist X eine Brownsche Bewegung bezüglich \mathbb{Q}_t . \square

Nun wird noch die Definition einer geometrischen Brownschen Bewegung angegeben, welche auch von einer Brownschen Bewegung abhängt. Ein Prozess dieser Art wird auch für die Simulation von Aktienkursen verwendet. In [Gla04] wird die Definition wie folgt hergeleitet.

Es sei W eine Brownsche Bewegung und X eine Brownsche Bewegung mit Drift, daher erfüllt X die stochastische Differentialgleichung

$$dX_t = \mu dt + \sigma dW_t.$$

Setzt man

$$S_t = S_0 \exp(X_t) \equiv f(X_t),$$

dann folgt durch Anwendung der Itô-Formel

$$\begin{aligned} dS_t &= f'(X_t) dX_t + \frac{1}{2} \sigma^2 f''(X_t) dt = S_0 \exp(X_t) \left(\mu dt + \sigma dW_t + \frac{1}{2} \sigma^2 dt \right) \\ &= S_t \left(\left(\mu + \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t \right). \end{aligned}$$

Definition 2.15. Ein Prozess $S = (S_t)_{0 \leq t \leq T}$ wird als geometrische Brownsche Bewegung bezeichnet, wenn er Lösung der stochastischen Differentialgleichung

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \quad (6)$$

ist, wobei W eine Brownsche Bewegung ist. μ wird als Drift-Parameter und σ wird als Volatilitätsparameter bezeichnet.

Mit der Itô-Formel kann man verifizieren, dass S eine geometrische Brownsche Bewegung mit Anfangswert S_0 ist, wenn

$$S_t = S_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right)$$

erfüllt. Etwas allgemeiner gilt

$$S_t = S_u \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) (t - u) + \sigma (W_t - W_u) \right).$$

Betrachtet man die geometrische Brownsche Bewegung unter dem zu \mathbb{P} äquivalenten Wahrscheinlichkeitsmaß \mathbb{Q} , bezüglich dem die hier verwendete Brownsche Bewegung mit Drift ein Martingal ist, dann ist die geometrische Brownsche Bewegung auch ein Martingal bezüglich \mathbb{Q} , nämlich das Exponentialmartingal der Brownschen Bewegung mit Drift.

2.6.2 Mehrdimensional

So wie im eindimensionalen Fall, kann man eine Brownsche Bewegung auch mehrdimensional definieren:

Definition 2.16. Ein Prozess $W = (W_t)_{0 \leq t \leq T}$ mit $W_t = (W_t^1, \dots, W_t^d)$ wird eine d -dimensionale Brownsche Bewegung genannt, wenn $W_0 = 0$, die Pfade stetig, die Zuwächse unabhängig und

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

für alle $0 \leq s < t \leq T$, und I ist die Einheitsmatrix.

Jedes W_t^i für $i = 1, \dots, d$ ist eine eindimensionale Brownsche Bewegung und W^i, W^j sind unabhängig für $i \neq j$.

Definition 2.17. Für einen Vektor $\mu \in \mathbb{R}^d$ und eine positiv definite oder semidefinite Matrix $\Sigma \in \mathbb{R}^{d \times d}$ wird ein Prozess X als d -dimensionale Brownsche Bewegung mit Drift bezeichnet, wenn X stetige Pfade und unabhängige Zuwächse hat mit

$$X_t - X_s \sim N((t - s)\mu, (t - s)\Sigma).$$

Der Anfangsvektor X_0 ist ein beliebiger Vektor. μ wird als Drift und Σ als Kovarianz bezeichnet.

Definition 2.18. Eine d -dimensionale geometrische Brownsche Bewegung kann durch ein System stochastischer Differentialgleichungen der Form

$$\frac{dS_t^i}{S_t^i} = \mu_i dt + \sigma_i dW_t^i \text{ für } i = 1, \dots, d \quad (7)$$

wobei W^i eine eindimensionale Brownsche Bewegung ist, definiert werden. Es gilt, dass W_t^i, W_t^j Korrelation ρ_{ij} haben.

2.7 Finanzmathematische Grundlagen

In diesem Kapitel werden kurz benötigte Definitionen und Sätze aus der Finanzmathematik angegeben, welche benötigt werden. Dabei orientiert man sich an dem Buch [FS11].

Es wird ein Marktmodell angenommen, in welchem $d+1$ Finanzinstrumente zu den Zeitpunkten $t = 0, 1, \dots, T$ bepreist werden. Der Preis des Finanzinstrumentes i zur Zeit t ist eine nicht negative Zufallsvariable S_t^i auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{B}, \mathbb{P})$. Der Zufallsvektor $S_t = (S_t^0, S_t^1, \dots, S_t^d)$ ist bezüglich einer σ -Algebra $\mathcal{F}_t \subset \mathcal{F}$ messbar. Es gilt

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_T.$$

also ist die Familie der σ -Algebren $(\mathcal{F}_t)_{t=0, \dots, T}$ eine Filtration. Es wird angenommen, dass die Filtration die üblichen Bedingungen erfüllt. Den Raum $(\Omega, \mathcal{B}, (\mathcal{F}_t)_{t=0, \dots, T}, \mathbb{P})$ nennt man filtrierten Wahrscheinlichkeitsraum.

Die Filtration beinhaltet alle Informationen, welche zu jenem Zeitpunkt beobachtbar sind.

Definition 2.19. Eine Handelsstrategie ist ein vorhersagbarer \mathbb{R}^{d+1} -wertiger Prozess

$$\xi = (\xi_t^0, \xi_t^1, \dots, \xi_t^d).$$

Der Wert ξ_t^1 beschreibt die Anzahl der Aktien des Finanzinstrumentes i , welche man in der Periode $t-1$ und t hält. Daher ist $\xi_t^i S_{t-1}^i$ der Betrag, welcher zum Zeitpunkt $t-1$ in das Finanzinstrument i investiert wurde. Damit ergibt sich der Wert des Portfolios zur Zeit $t-1$ durch

$$\xi_t S_{t-1} = \sum_{i=1}^d \xi_t^i S_{t-1}^i.$$

Zur Zeit t hat sich der Wert des Portfolios dann zu

$$\xi_t S_t = \sum_{i=1}^d \xi_t^i S_t^i$$

verändert.

Definition 2.20. Eine Handelsstrategie ξ heißt selbstfinanzierend, wenn

$$\xi_t S_t = \xi_{t+1} S_t$$

für $t = 1, \dots, T - 1$.

Was bedeutet, dass das Portfolio immer so arrangiert wird, dass der aktuelle Wert wieder ganz investiert wird. Es wird kein neues Geld investiert und es gibt auch keine Auszahlung.

Nimmt man nun an, dass

$$S_t^0 > 0 \text{ } \mathbb{P}\text{-fast sicher}$$

für alle t , zum Beispiel ein Bankkonto mit fixer Verzinsung, dann definiert man den diskontierten Preisprozess durch

$$X_t^i = \frac{S_t^i}{S_t^0}$$

für alle $t = 0, \dots, T$ und $i = 1, \dots, d$.

Der diskontierte Preisprozess, gibt die Änderung des Preises in Relation zum Preis des Finanzinstruments S^0 , welches auch als Numéraire bezeichnet wird, an. Damit kann man beobachten um wie viel die Investition in ein Finanzinstrument, welches ein Risiko aufweist, besser ist als das Geld nur auf ein Bankkonto zu geben.

Definition 2.21. Eine selbstfinanzierende Handelsstrategie wird Arbitrage genannt, wenn der Wertprozess der Handelsstrategie folgende Eigenschaften erfüllt

- $V_0 \leq 0$
- $V_T \geq 0$ \mathbb{P} -fast sicher
- $\mathbb{P}(V_T > 0) > 0$.

Bei einer Arbitrage wäre es möglich ohne Kosten einen risikolosen Gewinn zu machen. Ein Marktmodell in dem keine Arbitrage erlaubt wird, wird als arbitragefrei bezeichnet.

Es wird in dieser Arbeit von einem arbitragefreien Markt ausgegangen. Daher kann man durch die Anwendung des Fundamentalsatzes der Arbitrage-theorie davon ausgehen, dass es ein zu \mathbb{P} äquivalentes Martingalmaß oder auch risikoneutrales Maß \mathbb{Q} gibt. Das bedeutet, dass ein zu \mathbb{P} äquivalentes Wahrscheinlichkeitsmaß \mathbb{Q} existiert bezüglich dem der diskontierte Preisprozess X ein Martingal ist. Der Fundamentalsatz der Arbitrage-theorie kann wie folgt formuliert werden:

Satz 2.14. *Ein Marktmodell ist genau dann arbitragefrei, wenn zumindest ein äquivalentes Martingalmaß \mathbb{Q} zum realen Maß \mathbb{P} existiert.*

Beweis. Siehe [FS11] (Satz 5.16). \square

Da die beiden Maße äquivalent sind, ist die Dichte $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$ \mathbb{P} -fast sicher positiv und es gilt auch $\frac{d\mathbb{P}}{d\mathbb{Q}} = Z^{-1} > 0$ \mathbb{P} -fast sicher.

Des Weiteren seien die folgende Konstanten endlich

$$C_1^{\mathbb{P}} = \|D_T\|_{L^2(\mathbb{P})}, \quad C_2^{\mathbb{P}} = \left\| \frac{D_T}{D_1} \right\|_{L^2(\mathbb{P})},$$

$$C_1^{\mathbb{Q}} = \left\| \frac{1}{D_T} \right\|_{L^2(\mathbb{Q})}, \quad C_2^{\mathbb{Q}} = \left\| \frac{1}{D_1} \right\|_{L^2(\mathbb{Q})},$$

wobei mit L^1 und L^2 die in 2.2 erläuterten Banach- bzw. Hilberträume gemeint sind. Da $\frac{1}{D_t}$ ein \mathbb{Q} -Martingal ist, gilt $C_2^{\mathbb{Q}} \leq C_1^{\mathbb{Q}}$.

Mit dieser Annahme kann das folgende Lemma, welches auch in [CF17] angeführt wird, bewiesen werden, was bei der Fehleranalyse des replizierenden Portfolios von Nutzen sein wird.

Lemma 2.4. *Für eine Zufallsvariable X gilt*

$$\|X\|_{L^1(\mathbb{Q})} \leq C_1^{\mathbb{P}} \|X\|_{L^2(\mathbb{P})}, \quad \|X\|_{L^1(\mathbb{P})} \leq C_1^{\mathbb{Q}} \|X\|_{L^2(\mathbb{Q})}.$$

Also gilt $L^2(\mathbb{P}) \subset L^1(\mathbb{Q})$ und $L^2(\mathbb{Q}) \subset L^1(\mathbb{P})$. Wenn $X \in L^2(\mathbb{M})$, dann ist

$$\|E_{\mathbb{Q}}[X \mid \mathcal{F}_1]\|_{L^1(\mathbb{P})} \leq C_2^{\mathbb{M}} \|X\|_{L^2(\mathbb{M})}$$

Beweis. Die ersten beiden Behauptungen folgen analog aus der Cauchy-Schwarzsche-Ungleichung:

$$\|X\|_{L^1(\mathbb{Q})} = E_{\mathbb{Q}}[|X|] = E_{\mathbb{P}}[|X| D_T] = \langle |X|, D_T \rangle_{L^2(\mathbb{P})}$$

$$\stackrel{\text{Cauchy-Schwarz}}{\leq} \|D_T\|_{L^2(\mathbb{P})} \|X\|_{L^2(\mathbb{P})}$$

Für die zweite Behauptung gilt mit der Formel von Bayes

$$\|E_{\mathbb{Q}}[X \mid \mathcal{F}_1]\|_{L^1(\mathbb{P})} = \left\| \frac{E_{\mathbb{P}}[X D_T \mid \mathcal{F}_1]}{D_1} \right\|_{L^1(\mathbb{P})}$$

$$\leq \left\| X \frac{D_T}{D_1} \right\|_{L^1(\mathbb{P})} = \left\| X \frac{1}{D_1} \right\|_{L^1(\mathbb{Q})},$$

wenn man nun die Cauchy-Schwarzsche-Ungleichung auf den letzten und vorletzten Ausdruck anwendet, erhält man die Behauptung für $\mathbb{M} = \mathbb{Q}$ und \mathbb{P} . \square

In realen Finanzmärkten werden nicht nur Finanzinstrumente, wie zum Beispiel Aktien oder Anleihen gehandelt, sondern es werden eine Vielzahl von Wertpapieren gehandelt, deren Auszahlung von den zugrundeliegenden bereits erwähnten Finanzinstrumenten abhängen. Diese Art von Finanzinstrumenten werden im folgenden als Finanzderivate oder auch Optionen bezeichnet.

In den Beispielen im Kapitel 9, werden unter anderem auch verschiedene Finanzderivate betrachtet.

3 Modellrahmen

Es wird nun ein Finanzmarktmodell festgelegt, welches für die gesamte Arbeit gleich bleibt.

T sei der fixe Zeithorizont unseres Modells. Dieser wird durch die Laufzeit jener Verbindlichkeit festgelegt, dessen Auszahlungszeitpunkt zur Zeit der Berechnung am weitesten in der Zukunft liegt. Der Zeithorizont wird in den Risikohorizont $[0, t]$, welcher bei Solvency II und dem Schweizer Solvenztest ein Jahr also $t = 1$ beträgt, und in den Projektionshorizont $(t, T]$ aufgeteilt. Der Risikohorizont ist jener Zeitraum über den das Risiko gemessen werden soll.

Es sei $(\Omega, \mathcal{B}, \mathcal{F}_t, \mathbb{P})$ ein filtrierter Wahrscheinlichkeitsraum, wobei mit \mathbb{P} das reale Wahrscheinlichkeitsmaß bezeichnet wird, und $(\mathcal{F}_t)_{0 \leq t \leq T}$ ist eine Filtration, welche die üblichen Bedingungen erfüllt, dass heißt \mathcal{F}_0 enthält alle \mathbb{P} Nullmengen und $(\mathcal{F}_t)_{0 \leq t \leq T}$ ist rechtsstetig. Sie beschreibt den Informationsverlauf, dass bedeutet \mathcal{F}_t beinhaltet alle Informationen, welche zum Zeitpunkt t bekannt sind.

Da von einem arbitragefreien Markt ausgegangen wird, was bedeutet, dass bei einem Portfolio kein risikoloser Gewinn ohne Kosten aufgebaut werden kann, existiert nach dem Fundamentalsatz der Arbitrage-Theorie ein zu \mathbb{P} äquivalentes Martingalmaß \mathbb{Q} .

In dieser Arbeit wird der Numéraire im ersten Jahr als konstant angenommen, womit er keine Rolle bei den Berechnungen spielt, daher wird er nicht weiter beachtet.

Es sei $G = (G_t)_{0 \leq t \leq T}$ ein d -dimensionaler stochastischer Prozess. $t \mapsto G(\omega, t)$, $t \in [0, T]$ liegt im Funktionsraum $\mathbb{D}_d[0, T]$, welcher alle Funktionen $f : [0, T] \rightarrow \mathbb{R}^d$ enthält. G beschreibt, den Prozess der Finanzinstrumente, welche zur Replikation des Portfolios verwendet werden. Was bedeutet, dass die Basis, welche verwendet wird von G abhängt. In den Beispielen die betrachtet werden, ist G eine Brownsche Bewegung. Es wird auf Grund der Arbitragefreiheit angenommen, dass G ein Martingal bzgl. \mathbb{Q} ist. Da die Filtration \mathcal{F} den gesamten Informationsverlauf enthält ist G an \mathcal{F} adaptiert, was bedeutet, dass G_t \mathcal{F}_t -messbar ist für alle $0 \leq t \leq T$.

Weiter sei X die diskontierte Endauszahlung des Asset-Liability Portfolios zur Zeit T , welche in Thermen von G dargestellt werden soll. Es wird angenommen, dass $X \in L^2(\mathbb{M})$, was bedeutet, dass X quadratisch integrierbar ist, also existiert der Erwartungswert und die Varianz. Darüber hinaus wird davon ausgegangen, dass das Portfolio zum Zeitpunkt $t = 0$ unter \mathbb{Q} fair

bepreist wurde, also gilt $E_{\mathbb{Q}}[X] = 0$. Des Weiteren ist X \mathcal{F}_T -messbar.

Es ist möglich, dass die Auszahlung pfadabhängig ist, daher sei $A_t : \mathbb{D}_d[0, t] \rightarrow \mathbb{R}^{l_t}$ für $t \in (0, T]$ eine beliebige messbare Funktion, so dass $(A_t(G))$ ein l_t -dimensionaler Prozess ist, welcher alle relevanten Informationen für die Auszahlungsfunktion auf den Pfaden von $(G_s)_{0 \leq s \leq t}$ beinhaltet. Die Funktion $A_t(G)$ ist nicht eindeutig, es gibt einen gewissen Spielraum bei der Wahl der Funktion, was auch Auswirkungen auf die Dimension hat, worauf später in Kapitel 6.2.3 genauer eingegangen wird.

Man kann $X = g_T(A_T(G))$ für alle $\omega \in \Omega$ schreiben, wobei $g_T : \mathbb{R}^{l_T} \rightarrow \mathbb{R}$ Borel-messbar ist.

Dazu ein kurzes Beispiel um diese Definition etwas verständlicher zu machen.

Beispiel 3.1. *Betrachtet man eine Asiatische Put-Option, deren Auszahlung zum Zeitpunkt T , wie folgt definiert ist*

$$X = \left(K - \left(\prod_{i=0}^m G_{t_i} \right)^{\frac{1}{m}} \right)$$

wobei $0 = t_0 < t_1 < \dots < t_m = T$ und G ist ein einfaches Finanzinstrument, also zum Beispiel der Wertprozess einer Aktie.

Dann ist es möglich die Funktion als das Gesamte Produkt zu definieren

$$A_T(G) = \prod_{i=0}^m G_{t_i},$$

oder als die einzelnen G_{t_i} für $i = 0, \dots, m$. Im ersten Fall ist die Dimension $L_T = 1$ und im zweiten Fall ist $L_T = m + 1$.

4 Die Risikomaße Value at Risk und Expected Shortfall

Das Solvenzkapital eines Versicherungsunternehmens wird mit Risikomaßen berechnet, im Solvency II wird der Value at Risk und im SST der Expected Shortfall als Standardmaß angegeben. Beide Risikomaße werden über die Gewinn- beziehungsweise Verlustverteilung eines Portfolios berechnet, wobei für ein Portfolio der Gewinn zum Zeitpunkt T durch die Zufallsvariable $V_T - V_0$ definiert ist, dabei ist $V = (V_t)_{0 \leq t \leq T}$ der Wert des Portfolios zum Zeitpunkt t . Der Verlust ist als negativer Gewinn definiert, also gilt $-(V_T - V_0)$ ist der Verlust. Hier werden die Definition über die Gewinnverteilungen angegeben.

Dieses Kapitel orientiert sich an [Kre18], [FS11] und [KW12].

4.1 Definition der Risikomaße

Risiko wird in [GKW18] als

Kennzeichnung der Eventualität, dass mit einer (ggf. niedrigen, ggf. auch unbekannt) Wahrscheinlichkeit ein (ggf. hoher, ggf. in seinem Ausmaß unbekannter) Schaden bei einer (wirtschaftlichen) Entscheidung eintritt oder ein erwarteter Vorteil ausbleiben kann

definiert.

Um das Risiko eines Portfolios zu bewerten, müsste man die Verteilung der Gewinne kennen. Da dies meistens nicht der Fall ist, bewertet man mit Hilfe von Risikomaßen, welche ganz allgemein wie folgt definiert sind.

Definition 4.1. *Ein Risikomaß ρ ist eine Abbildung, welche die Zufallsvariablen in die reellen Zahlen abbildet, das heißt es ordnet jeder Zufallsvariable X einen Wert $\rho(X) \in \mathbb{R}$ zu.*

Es gibt mittlerweile eine große Anzahl an Risikomaßen, als klassisches Maß wurde immer die Standardabweichung betrachtet. Diese bewertet jedoch Gewinne und Verluste gleich. Der Expected Shortfall und Value at Risk werden über das Quantil der Gewinnverteilung definiert. Sie betrachten nur die kleinsten Gewinne beziehungsweise was äquivalent dazu ist die größten Verluste.

Da der Value at Risk und der Expected Shortfall die Standardmaße in Solvency II und dem Swiss Solvenz Test sind, werden hier nur diese zwei genauer betrachtet.

Der Value at Risk zum Level α kann durch den Zusammenhang der oberen

und unteren Quantil-Funktionen, wie in [FS11], auf mehrere Arten definiert werden:

Definition 4.2. Sei $\alpha \in (0, 1)$ beliebig. Für ein Finanzinstrument X definiert man den Value at Risk zum Level α als

$$VaR_\alpha(X) := -q_X^+(1 - \alpha) = q_{-X}^-(\alpha) = \inf \{m \mid \mathbb{P}(X + m < 0) \leq 1 - \alpha\} \quad (8)$$

Der Expected Shortfall wird in [Kre18] wie folgt definiert:

Definition 4.3. Sei $\alpha \in (0, 1)$ beliebig und X ein Finanzinstrument. Es gelte $E[X^-] < \infty$, dann wird der Expected Shortfall zum Level α von X definiert durch

$$\begin{aligned} ES_\alpha(X) &= \\ &= -\frac{1}{1 - \alpha} \left(E \left[X \mathbf{1}_{\{X \leq q_X^-(1 - \alpha)\}} \right] + q_X^-(1 - \alpha) (1 - \alpha - \mathbb{P}(X \leq q_X^-(1 - \alpha))) \right). \end{aligned} \quad (9)$$

Wird angenommen, dass $F(q_X^-(1 - \alpha)) = \mathbb{P}(X \leq q_X^-(1 - \alpha)) = 1 - \alpha$, dann ist

$$\begin{aligned} ES_\alpha(X) &= -\frac{1}{1 - \alpha} E \left[X \mathbf{1}_{\{x \leq q_X^-(1 - \alpha)\}} \right] = E \left[X \mid X \leq q_X^-(1 - \alpha) \right] \\ &= \frac{1}{1 - \alpha} \int_0^{q_X^-(1 - \alpha)} x dF_X(x) \end{aligned}$$

Gilt darüber hinaus $q_X^+(1 - \alpha) = q_X^-(1 - \alpha)$, dann kann der Expected Shortfall in Abhängigkeit des Value at Risks dargestellt werden:

$$ES_\alpha(X) = -E \left[X \mid X \leq -VaR_\alpha(X) \right].$$

Am Anschaulichsten kann man den Unterschied dieser zwei Definitionen darstellen, wenn man annimmt, dass 1000 Stichproben simuliert wurden. Der Value at Risk zum Level 99.5 wäre dann der größte der 50 kleinsten Realisationen, während der Expected Shortfall der Durchschnitt dieser 50 kleinsten Stichproben wäre.

Ein Nachteil des Value at Risk ist, dass über die Höhe der extremen Verluste nichts ausgesagt werden kann, was den Anreiz geben könnte Portfolios zu bilden, welche vielleicht anfällig sind für extreme Ereignisse am Rand der Verteilung.

4.2 Kohärenz

Eine weitere wichtige Definition, welche in [Kre18] wie unten angegeben ist, in Bezug auf Risikomaße ist die Kohärenz. Ein Risikomaß, welches die folgenden vier Eigenschaften erfüllt, wird als kohärent bezeichnet.

Definition 4.4. *Man nennt ein Risikomaß ρ kohärent, wenn es für alle Risiken X und Y , für alle $a \in \mathbb{R}$ und für alle $t \geq 0$ folgende Eigenschaften erfüllt:*

1. Für $X \geq 0$ gilt $\rho(X) \leq 0$ (Monotonie).
2. $\rho(X + a) = \rho(X) - a$ (Translationsinvarianz)
3. $\rho(tX) = t\rho(X)$ (positive Homogenität)
4. $\rho(X + Y) \leq \rho(X) + \rho(Y)$ (Subadditivität)

Aus der Monotonie folgt, dass für eine positive Auszahlung X kein Sicherheitskapital bereit gestellt werden muss beziehungsweise, wenn man zwei unterschiedliche Finanzinstrumente betrachtet muss für jenes mehr Sicherheitskapital bereitgestellt werden, für welches der kleinere Gewinn prognostiziert wird.

Aus der Translationsinvarianz kann gefolgert werden, dass, wenn die Auszahlung um eine Konstante $a \in \mathbb{R}$ verkleinert oder vergrößert wird, sich auch das Sicherheitskapital um diese Konstante ändert. Diese Eigenschaft wird von der Standardabweichung nicht erfüllt.

Die positive Homogenität besagt, dass betrachtet man ein positives Vielfaches der Auszahlung, dann ändert sich auch das Sicherheitskapital um dieses Vielfache. Investiert man zum Beispiel zwei Einheiten anstelle nur einer in eine risikoreiche Aktie, dann ist es auch sinnvoll, wenn sich das Sicherheitskapital verdoppelt, welches man dafür zurückhalten muss.

Die am meisten umstrittene Eigenschaft der Kohärenz ist die Subadditivität. Welche besagt, dass das Sicherheitskapital, welches benötigt wird um zwei Auszahlungen gemeinsam abzudecken, kleiner ist als jenes, welches man benötigt um beide Auszahlungen getrennt von einander abzudecken. Was heißt, dass die Diversifikation das Risiko verringert. Jedoch soll die Subadditivität eines Risikomaßes nicht dafür missbraucht werden, dass durch eine geeignete Aufteilung des Portfolios das Sicherheitskapital klein gerechnet wird.

Es werden nun der Value at Risk und der Expected Shortfall in Bezug auf diese vier Eigenschaften betrachtet.

4.2.1 Value at Risk

1. Monotonie:

Sei $X \leq Y$, dann ist auch $X + m \leq Y + m$ für alle $m \in \mathbb{R}$. Daraus

folgt $\mathbb{P}[X + m < 0] \geq \mathbb{P}[Y + m < 0]$. Daher gilt auch $VaR_\alpha(X) \geq VaR_\alpha(Y)$.

2. Translationsinvarianz: Sei $a \in \mathbb{R}$, dann gilt

$$\begin{aligned} VaR_\alpha(X + a) &= \inf \{m \in \mathbb{R} \mid \mathbb{P}(X + a + m < 0) \leq 1 - \alpha\} \\ &\stackrel{l=m+a}{=} \inf \{l \in \mathbb{R} \mid \mathbb{P}(X + l < 0) \leq 1 - \alpha\} - a = VaR_\alpha(X) - a \end{aligned}$$

3. Positive Homogenität: Für $t \geq 0$ gilt

$$\begin{aligned} VaR_\alpha(tX) &= \inf \{m \in \mathbb{R} \mid \mathbb{P}(tX + m < 0) \leq 1 - \alpha\} \\ &= \inf \left\{ m \in \mathbb{R} \mid \mathbb{P}\left(X + \frac{m}{t} < 0\right) \leq 1 - \alpha \right\} \\ &\stackrel{l=\frac{m}{t}}{=} t \inf \{l \in \mathbb{R} \mid \mathbb{P}(X + l < 0) \leq 1 - \alpha\} \\ &= tVaR_\alpha(X) \end{aligned}$$

4. Subadditivität:

Der Value at Risk ist im allgemeinen nicht subadditiv, das wird anhand eines Gegenbeispiels gezeigt.

Sei $(\Omega, \mathcal{B}, \mathbb{P})$ ein Wahrscheinlichkeitsraum mit $\Omega = \{\omega_1, \omega_2\}$. \mathcal{B} ist die Menge aller Teilmengen von Ω und \mathbb{P} ist ein Wahrscheinlichkeitsmaß, welches wie folgt definiert wird:

$$\mathbb{P}(\{\omega_1\}) = p \text{ und } \mathbb{P}(\{\omega_2\}) = 1 - p.$$

Es seien X_i für $i = 1, 2$ zwei unabhängige identisch verteilte Zufallsvariablen auf $(\Omega, \mathcal{B}, \mathbb{P})$, welche wie folgt definiert werden:

$$X_i(\omega_j) = \begin{cases} 1, & \text{für } j = 1 \\ -2, & \text{für } j = 2 \end{cases}$$

Weiters wird angenommen, dass $p \geq \alpha$ ist. Dann ist $VaR_\alpha(X_i) = -1$ für $i = 1, 2$, denn

$$\mathbb{P}(X_i + m < 0) = \begin{cases} 1, & \text{wenn } m < -1 \\ 1 - p \leq 1 - \alpha, & \text{wenn } m \geq -1 \end{cases}$$

Daher ist also $VaR_\alpha(X_1) + VaR_\alpha(X_2) = -2$. Nun betrachtet man die Zufallsvariable $X_1 + X_2$, welche wie folgt definiert wird:

$$X_1(\omega_i) + X_2(\omega_j) = \begin{cases} 2, & \text{für } i, j = 1 \\ -1, & \text{für } i = 1, j = 2 \text{ oder } i = 2, j = 1 \\ -4, & \text{für } i, j = 2 \end{cases}$$

Da der Value at Risk bestimmt werden soll, sieht man sich $\mathbb{P}(X_1 + X_2 + m < 0)$ an:

$$\mathbb{P}(X_1 + X_2 + m < 0) = \begin{cases} 1, & \text{wenn } m < -2 \\ 1 - p^2, & \text{wenn } -2 \leq m < 1 \\ (1 - p)^2, & \text{wenn } m \geq 1 \end{cases}$$

Gilt $1 - p \leq 1 - \alpha < 1 - p^2$, was zum Beispiel bei $\alpha = 0.99$ und $p = 0.991$ der Fall wäre, ist

$$VaR_\alpha(X_1 + X_2) = 1 \geq -2 = VaR_\alpha(X_1) + VaR_\alpha(X_2)$$

Daher er die Subadditivität nicht erfüllt, ist der Value at Risk nicht kohärent. Das nicht erfüllen dieser Eigenschaft bedeutet, dass es möglich ist, dass der Value at Risk ein höheres Eigenmittel prognostiziert um das gesamte Portfolio abzudecken, als wenn man den benötigten Eigenmittelbedarf von Teilportfolios summiert. Also fördert der Value at Risk die Diversifikation von Risiken im Allgemeinen nicht.

4.2.2 Expected Shortfall

Im Gegensatz dazu erfüllt der Expected Shortfall alle vier Eigenschaften, demnach ist er ein kohärentes Risikomaß.

Um die Kohärenz des Expected Shortfall zu beweisen benötigt es einige Vorarbeit, welche aus dem Buch [Kre18] entnommen wurde.

Lemma 4.1. *Für jedes $\alpha \in (0, 1)$ gilt*

$$ES_\alpha(X) = -\frac{1}{1 - \alpha} E \left[X \mathbf{1}_{\{x \leq q_X^-(1 - \alpha)\}}^{(\alpha)} \right] \quad (10)$$

wobei

$$\mathbf{1}_{\{x \leq X\}}^{(\alpha)} = \begin{cases} \mathbf{1}_{\{x \leq X\}}, & \text{für } \mathbb{P}(X = x) = 0, \\ \mathbf{1}_{\{x \leq X\}} + \frac{1 - \alpha - \mathbb{P}(X \leq x)}{\mathbb{P}(X = x)} \mathbf{1}_{\{x = X\}}, & \text{für } \mathbb{P}(X = x) > 0 \end{cases} \quad (11)$$

ist. Weiters gilt

$$\mathbf{1}_{\{x \leq q_X^-(1 - \alpha)\}}^{(\alpha)} \in [0, 1] \quad (12)$$

und

$$E \left[\mathbf{1}_{\{x \leq q_X^-(1 - \alpha)\}}^{(\alpha)} \right] = 1 - \alpha. \quad (13)$$

Beweis. 1. Fall: $\mathbb{P}(X = q_X^-(1 - \alpha)) = 0$

In diesem Fall ist F an der Stelle $q_X^-(1 - \alpha)$ stetig, und es gilt $F(q_X^-(1 - \alpha)) =$

$1 - \alpha$. Nach (11) gilt $\mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}^{(\alpha)} = \mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}$, damit erhält man (10), die beiden anderen Behauptungen folgen da $E \left[\mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}} \right] = F(q_X^-(1 - \alpha)) = 1 - \alpha$.

2.Fall: $\mathbb{P}(X = q_X^-(1 - \alpha)) > 0$

In diesem Fall besitzt F an der Stelle $q_X^-(1 - \alpha)$ eine Sprungstelle mit mit Sprunghöhe $\mathbb{P}(X = q_X^-(1 - \alpha))$. Es wird nun verwendet, dass

$$\mathbb{P}(X < q_X^-(1 - \alpha)) + \mathbb{P}(X = q_X^-(1 - \alpha)) = \mathbb{P}(X \leq q_X^-(1 - \alpha)) = F(q_X^-(1 - \alpha))$$

und, dass für jede streng monoton wachsende Folge $(x_n)_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} x_n = q_X^-(1 - \alpha)$

$$\mathbb{P}(X < q_X^-(1 - \alpha)) = \lim_{n \rightarrow \infty} F(x_n)$$

gilt. Aus der Definition des unteren $(1 - \alpha)$ -Quantils folgt $F(x_n) < 1 - \alpha$ für jedes n und daher

$$\mathbb{P}(X < q_X^-(1 - \alpha)) = \lim_{n \rightarrow \infty} F(x_n) \leq 1 - \alpha.$$

Zusammen mit $F(q_X^-(1 - \alpha)) \geq 1 - \alpha$, was in Lemma 2.2 gezeigt wurde, folgt

$$\mathbb{P}(X = x) = F(q_X^-(1-\alpha)) - \mathbb{P}(X < q_X^-(1-\alpha)) \geq F(q_X^-(1-\alpha)) - (1-\alpha). \quad (14)$$

Also wird (11) zu

$$\mathbf{1}_{\{x \leq X\}}^{(\alpha)} = \begin{cases} 1, & \text{für } x < q_X^-(1 - \alpha) \\ 1 - \frac{F(q_X^-(1-\alpha)) - (1-\alpha)}{\mathbb{P}(X = q_X^-(1-\alpha))}, & \text{für } x = q_X^-(1 - \alpha) \\ 0, & \text{für } x > q_X^-(1 - \alpha). \end{cases} \quad (15)$$

Aus (14) und (15) folgt die Behauptung (12). Die anderen zwei Resultate kann man ganz einfach nachrechnen. \square

Mit diesem Lemma ist es und nun möglich die Kohärenz des Expected Shortfalls nachzuweisen.

Satz 4.1. *Sei $\alpha \in (0, 1)$ fest gewählt. Dann ist der Expected Shortfall $ES_\alpha : \Omega \rightarrow \mathbb{R}$ ein kohärentes Risikomaß.*

Beweis. 1. Monotonie:

Für $X \geq 0$ gilt $E \left[X \mathbf{1}_{\{X \leq q_X^-(1-\alpha)\}}^{(\alpha)} \right] \geq 0$ wegen (12). Zusammen mit (10) folgt $ES_\alpha \leq 0$, also die Monotonie.

2. Positive Homogenität:

Für $\lambda > 0$ gilt nach Korollar 2.1

$$q_{\lambda X}^-(1 - \alpha) = \lambda q_X^-(1 - \alpha)$$

also

$$\mathbb{P}(\lambda X \leq q_{\lambda X}^-(1-\alpha)) = \mathbb{P}(X \leq q_X^-(1-\alpha)).$$

Die positive Homogenität des Expected Shortfalls folgt aus der Darstellung aus der Definition.

3. Translationsinvarianz:

Sei $a \in \mathbb{R}$ beliebig. Dann gilt nach Korollar 2.1

$$q_{X+a}^-(1-\alpha) = q_X^-(1-\alpha) + a$$

und daher auch

$$\begin{aligned} E \left[(X+a) \mathbf{1}_{\{X+a \leq q_{X+a}^-(1-\alpha)\}} \right] &= E \left[X \mathbf{1}_{\{X \leq q_X^-(1-\alpha)\}} \right] + a E \left[\mathbf{1}_{\{X \leq q_X^-(1-\alpha)\}} \right] \\ &= E \left[X \mathbf{1}_{\{X \leq q_X^-(1-\alpha)\}} \right] + a \mathbb{P}(X \leq q_X^-(1-\alpha)). \end{aligned}$$

Desweiteren gilt

$$\begin{aligned} & q_{X+a}^-(1-\alpha)(1-\alpha - \mathbb{P}(X+a \leq q_{X+a}^-(1-\alpha))) \\ &= (q_X^-(1-\alpha) + a)(1-\alpha - \mathbb{P}(X \leq q_X^-(1-\alpha))) \\ &= q_X^-(1-\alpha)(1-\alpha - \mathbb{P}(X \leq q_X^-(1-\alpha))) + (1-\alpha)a - a \mathbb{P}(X \leq q_X^-(1-\alpha)). \end{aligned}$$

Setzt man diese Ergebnisse in die Definition ein erhält man die Translationsinvarianz des Expected Shortfalls.

4. Subadditivität:

Es sei $Z = X + Y$. Mit (12) und (15) gilt

$$\mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - \mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}^{(\alpha)} \begin{cases} \leq 0, & \text{für } X < q_X^-(1-\alpha) \\ \geq 0, & \text{für } X > q_X^-(1-\alpha) \end{cases} \quad (16)$$

und daraus folgt

$$X(-q_X^-(1-\alpha))(\mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - \mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}^{(\alpha)}) \geq 0.$$

Was zusammen mit (10) und (13) folgendes ergibt:

$$\begin{aligned} & \alpha(ES_\alpha(X) + ES_\alpha(Y) - ES_\alpha(Z)) \\ &= E \left[Z \mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - X \mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}^{(\alpha)} - Y \mathbf{1}_{\{y \leq q_Y^-(1-\alpha)\}}^{(\alpha)} \right] \\ &= E \left[X(\mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - \mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}^{(\alpha)}) + Y(\mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - \mathbf{1}_{\{y \leq q_Y^-(1-\alpha)\}}^{(\alpha)}) \right] \\ &\geq q_X^-(1-\alpha) E \left[\mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - \mathbf{1}_{\{x \leq q_X^-(1-\alpha)\}}^{(\alpha)} \right] + \\ &\quad + q_Y^-(1-\alpha) E \left[\mathbf{1}_{\{z \leq q_Z^-(1-\alpha)\}}^{(\alpha)} - \mathbf{1}_{\{y \leq q_Y^-(1-\alpha)\}}^{(\alpha)} \right] \\ &= q_X^-(1-\alpha)((1-\alpha) - (1-\alpha)) + q_Y^-(1-\alpha)((1-\alpha) - (1-\alpha)) = 0 \end{aligned}$$

□

4.3 Weitere Eigenschaften des Value at Risk und Expected Shortfall

Es werden nun einige Eigenschaften der zwei Risikomaße angeführt, welche im weiteren Verlauf der Arbeit benötigt werden.

Konvergenz und Monotonie in Bezug auf das Konvidenzniveau: Sowohl der Value at Risk als auch der Expected Shortfall sind monoton bezüglich α . Jedoch ist die Abbildung $\alpha \mapsto VaR_\alpha(X)$ für ein X fix nicht stetig, was zu großen Unterschieden der Eigenkapitalanforderung bei nur kleinen Änderungen des Konvidenzniveaus führen kann. Im Gegensatz dazu wird in [AT02] gezeigt, dass der Expected Shortfall für ein fixes X als Abbildung bzgl. α , also $\alpha \mapsto ES_\alpha(X)$ stetig ist.

Konvergenzeigenschaften: Für den Value at Risk gilt, folgende Konvergenzeigenschaft, wenn die Verteilungsfunktion, stetig und streng monoton steigend in α ist.

Satz 4.2. *Konvergiert eine Folge von Zufallsvariablen X_n schwach gegen eine Zufallsvariable X , deren rechtes und linkes $(1 - \alpha)$ -Quantil gleich sind, also $q_X^-(1 - \alpha) = q_X^+(1 - \alpha)$, dann gilt:*

$$\lim_{n \rightarrow \infty} VaR_\alpha(X_n) = VaR_\alpha(X) \quad (17)$$

Beweis. Siehe [CDS10] (Proposition 3.5) □

Für den Expected Shortfall gilt, nach [KW12]:

Satz 4.3. *Sei Y eine positive, integrierbare Funktion und $(X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen mit $|X_k| \leq Y$ fast sicher, die fast sicher punktweise gegen die Zufallsvariable X konvergieren. Dann gilt*

$$ES_\alpha(X_n) \rightarrow ES_\alpha(X)$$

Beweis. Siehe [KW12] (Proposition 2.4) □

Da die beiden Risikomaße meistens nicht direkt berechnet werden können, müssen sie simuliert werden, dabei kann folgender Satz, zur Vereinfachung der Berechnung, verwendet werden. Dieser Satz ist auch die Grundlage für die Simulation der Risikomaße in den Anwendungsbeispielen in Kapitel 9.

Satz 4.4. Sei X eine Zufallsvariable, welche

$$\begin{cases} q_X^+(1 - \alpha) = q_X^-(1 - \alpha), & \text{wenn } \rho = \text{VaR}_\alpha(X) \\ X \in L^1(\mathbb{P}), & \text{wenn } \rho = \text{ES}_\alpha(X) \end{cases}$$

erfüllt. Sei $(X^{(j)}, \frac{d\mathbb{P}}{d\mathbb{M}^{(j)}})$ für $j \geq 1$ eine i.i.d. Folge von Zufallsvariablen mit der selben \mathbb{M} -Verteilung wie $(X, \frac{d\mathbb{P}}{d\mathbb{M}})$. Für ein $n > 1$ definiere man die Gewichte

$$w^{(j)} = \frac{\frac{d\mathbb{P}}{d\mathbb{M}^{(j)}}}{\sum_{k=1}^n \frac{d\mathbb{P}}{d\mathbb{M}^{(k)}}} \left(= \frac{1}{n} \text{ für } \mathbb{M} = \mathbb{P} \right).$$

Mit denen ist die empirische \mathbb{P} -Verteilung von $X^{(1)}, \dots, X^{(n)}$

$$\hat{\mu}_n(x) = \sum_{j=1}^n w^{(j)} \mathbf{1}_{\{X^{(j)} \leq x\}}.$$

Dann gilt $\rho(\hat{\mu}_n) \rightarrow \rho(X)$ fast sicher für $n \rightarrow \infty$.

Beweis. Siehe [CF17] (Satz 4.4) □

Lipschitzstetigkeit: Wie in [KW12] gezeigt wird, kann der Expected Shortfall in folgender Weise geschrieben werden:

Satz 4.5. Es sei

$$\mathcal{Q}_\alpha = \left\{ \mathbb{Q} \mid \mathbb{Q} \text{ ist ein Wahrscheinlichkeitsmaß mit } \mathbb{Q} \ll \mathbb{P} \text{ und } \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{1 - \alpha} \right\}.$$

Dann gilt für eine integrierbare Zufallsvariable X

$$\text{ES}_\alpha(X) = \sup_{\mathbb{Q} \in \mathcal{Q}_\alpha} \{E_{\mathbb{Q}}[X]\}. \quad (18)$$

Mit dieser Darstellung kann gezeigt werden, dass der Expected Shortfall Lipschitzstetig auf $L^1(\mathbb{P})$.

Satz 4.6. Für zwei Finanzinstrumente X und Y gilt

$$|\text{ES}_\alpha(X) - \text{ES}_\alpha(Y)| \leq \frac{1}{1 - \alpha} \|X - Y\|_{L^1(\mathbb{P})}. \quad (19)$$

Beweis. Wegen der Subadditivität des Expected Shortfalls gilt

$$\text{ES}_\alpha(X) \leq \text{ES}_\alpha(X - Y) + \text{ES}_\alpha(Y)$$

und mit der Darstellung (18) folgt

$$\text{ES}_\alpha(X) - \text{ES}_\alpha(Y) \leq \text{ES}_\alpha(X - Y) \leq |\text{ES}_\alpha(X - Y)|$$

$$= \left| \sup_{\mathbb{Q} \in \mathcal{Q}_\alpha} E_{\mathbb{P}} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} (X - Y) \right] \right| \leq \frac{1}{1 - \alpha} E_{\mathbb{P}} [|X - Y|].$$

Vertauscht man X und Y erhält man

$$ES_\alpha(Y) - ES_\alpha(X) \leq \frac{1}{1 - \alpha} E_{\mathbb{P}} [|X - Y|]$$

und damit die Behauptung. \square

Der Value at Risk ist nicht Lipschitzstetig, jedoch wurde in der Arbeit [Nat18] eine ähnliche Eigenschaft mit Hilfe der Prokhorov-Metrik hergeleitet, welche jedoch mehrere Voraussetzungen benötigt. Die Prokhorov-Metrik wird in [Hub04] wie folgt definiert:

Definition 4.5. Sei \mathcal{M} der Raum aller Wahrscheinlichkeitsmaße auf Ω . Die Prokhorov-Metrik ist für $F, G \in \mathcal{M}$ wie folgt definiert:

$$d_p(F, G) = \inf \{ \epsilon > 0 \mid F(A) \leq G(A^\epsilon) + \epsilon, \text{ für alle } A \in \mathcal{B} \},$$

wobei A^ϵ die abgeschlossene ϵ -Umgebung von $A \in \mathcal{B}$ ist,

$$A^\epsilon = \left\{ x \in \Omega \mid \inf_{y \in A} d(x, y) \leq \epsilon \right\}.$$

Lemma 4.2. Sei $\alpha \in (0, 1)$ und X eine beliebige Zufallsvariable mit Verteilungsfunktion F_X , die an der Stelle $VaR_\alpha(X)$ differenzierbar ist, mit

$$F'(VaR_\alpha(X)) > \gamma \tag{20}$$

für $\gamma > 0$. Dann existiert eine Konstante $\delta > 0$, so dass für eine Zufallsvariable Y mit $d_p(X, Y) < \delta$

$$|VaR_\alpha(X) - VaR_\alpha(Y)| \leq \left(1 + \frac{1}{\gamma} \right) d_p(X, Y)$$

gilt, wobei d_p die Prokhorov-Metrik bezeichnet.

Beweis. Siehe [Nat18] (Lemma 4.2.1) \square

Wenn die Voraussetzungen des Lemmas 4.2 erfüllt sind, wird in [Nat18] gezeigt, dass folgende „Lipschitz Eigenschaft“ für den Value at Risk hält.

Proposition 4.1. Für den Einjahresgewinn L und $\alpha \in (0, 1)$ seien die Voraussetzungen des Lemmas 4.2 erfüllt. Dann existiert eine Konstante $\delta > 0$, so dass für alle Y mit $d_p(L, Y) < \delta$

$$|VaR_\alpha(L) - VaR_\alpha(Y)| \leq \left(1 + \frac{1}{\gamma} \right) (E_{\mathbb{P}} [|X - Y|])^{\frac{1}{2}}$$

gilt.

Verteilungsinvarianz:

Definition 4.6. Ein Risikomaß ρ wird verteilungsinvariant genannt, wenn $\rho(X) = \rho(Y)$ für zwei Zufallsvariablen X, Y , welche die gleiche Verteilung unter \mathbb{P} haben.

Sowohl der Expected Shortfall als auch der Value at Risk sind verteilungsinvariant.

Definition 4.7. Seien μ und ν zwei beliebige Wahrscheinlichkeitsmaße auf \mathbb{R} . Man sagt, μ dominiert ν stochastisch, wenn

$$\int f d\mu \geq \int f d\nu$$

für alle beschränkten monoton wachsenden Funktionen f gilt.

Bemerkung 4.1. Es wird in [FS11] gezeigt, dass ein verteilungsinvariantes Risikomaß ρ , welches translationsinvariant und monoton ist, monoton in Bezug auf die stochastische Ordnung ist, was bedeutet, wenn ν durch μ stochastisch dominiert wird, dann gilt

$$\rho(X_\mu) \leq \rho(X_\nu),$$

für X_μ und X_ν zwei Zufallsvariablen mit Verteilung μ beziehungsweise ν .

5 Monte Carlo Simulation

In diesem Kapitel wird die Monte Carlo (MC) Methode, welche eine stochastische Simulation ist, erläutert. Zuerst wird die klassische Variante beschrieben, danach werden einige Vorgehensweisen besprochen, welche verwendet werden können, um die rechnerische Effizienz dieser Methode zu verbessern. Der Inhalt dieses Kapitels bezieht sich auf die Bücher [Gla04], [ABLM13] und [GNR12].

Die Monte Carlo Methode wurde im Manhattan Projekt in den 1940ern entwickelt. Der Name bezieht sich auf die Zufälligkeit, welche in dieser Methode involviert ist, und findet seinen Ursprung im Namen von Monacos Verwaltungsgebiet Monte Carlo mit seinen Casinos.

Monte Carlo Simulation kommt dann zum Einsatz, wenn ein Problem analytisch nicht lösbar, beziehungsweise die analytische Lösung sehr aufwändig wäre. Sie ist mittlerweile ein sehr wichtiges Instrument in der Finanzmathematik und wird für die Bepreisung von Derivaten und im Risikomanagement genutzt.

Monte Carlo ist eine Methode der stochastischen Simulation, wobei man bei dieser Art im Allgemeinen unabhängige Stichproben einer Zufallsvariable Z generiert. Hat man eine Menge von Stichproben der Größe n , wobei n groß genug gewählt werden muss, erstmals simuliert, kann diese verwendet werden um einen Wert, welcher von der Verteilung von Z abhängt, zu schätzen (z.B. den Erwartungswert, die Varianz oder die Quantile).

Es stellt sich die Frage, wann so ein Schätzer als guter Schätzer zu betrachten ist. Die folgenden zwei Eigenschaften sollte ein Schätzer erfüllen. Er sollte erwartungstreu sein:

Definition 5.1. Ein Schätzer $\hat{\theta}_n$ für einen Parameter θ , welcher aus einer Stichprobe der Größe n berechnet wird, heißt erwartungstreu, wenn

$$E[\hat{\theta}_n] = E[\theta].$$

Weiter sollte eine Folge von Schätzern konsistent sein:

Definition 5.2. Es sei $(\hat{\theta}_n)_{n \geq 1}$ eine Folge von Schätzern für einen Parameter θ . Die Folge von Schätzern ist konsistent, wenn $\hat{\theta}_n$ in Wahrscheinlichkeit gegen θ konvergiert. Man sagt sie ist stark konsistent, wenn die Konvergenz mit Wahrscheinlichkeit 1 hält, das heißt, wenn die Konvergenz fast sicher ist.

Die Konsistenz eines Schätzers bedeutet, dass bei wachsender Größe der Stichprobe die Genauigkeit des Schätzers zunimmt.

Die Genauigkeit der Schätzung wird durch ein Konfidenzintervall um die Schätzung herum reflektiert. Es wird hier nur eine kurze Erklärung dieses Begriffs angegeben, welcher sich an [LW06] hält. Ein Konfidenzintervall zum Konfidenzniveau $1 - p$ ist ein Intervall für welches folgendes gilt:

Es sei F_θ , mit $\theta \in \Theta$ die vermutete Verteilungsfunktion von X_1, \dots, X_n , welche unabhängig und identisch verteilt sind. Weiter seien $U = U(X_1, \dots, X_n)$ und $O = O(X_1, \dots, X_n)$ Zufallsvariablen, mit O und U zwei Abbildungen von $\mathbb{R}^n \rightarrow \Theta$. Ein Konfidenzintervall $[U(X_1, \dots, X_n), O(X_1, \dots, X_n)]$ zum Konfidenzniveau $1 - p$ (mit vorgegebenen $p \in (0, 1)$) wird definiert durch

$$\mathbb{P}(\theta \in [U(X_1, \dots, X_n), O(X_1, \dots, X_n)]) \geq p$$

für alle $\theta \in \Theta$.

5.1 Klassische Monte Carlo Simulation

Es wird zuerst die Schätzung des Erwartungswertes ganz allgemein betrachtet:

Nimmt man an, dass eine Stichprobe (z_1, \dots, z_n) der Größe n für eine s -dimensionale Zufallsvariable Z generiert wurde, dann kann $\theta = E[g(Z)]$ durch

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n g(z_i)$$

geschätzt werden. $\hat{\theta}_n$ ist ein erwartungstreuer Schätzer, da

$$E[\hat{\theta}_n] = \frac{1}{n} \sum_{i=1}^n E[g(z_i)] = E[g(Z)].$$

Nimmt man an, dass sowohl der Erwartungswert als auch die Varianz von $g(Z)$ existiert, folgt mit dem starken Gesetz der großen Zahlen

$$\lim_{n \rightarrow \infty} \hat{\theta}_n = \theta = E[g(Z)] \text{ fast sicher.}$$

Ist $Var[g(Z)] = \sigma^2 < \infty$, dann gilt

$$Var[\hat{\theta}_n] = \frac{1}{n^2} \sum_{i=1}^n Var[g(z_i)] = \frac{\sigma^2}{n},$$

und mit dem zentralen Grenzwertsatz folgt

$$\frac{\hat{\theta}_n - \theta}{\frac{\sigma}{\sqrt{n}}} \underset{\text{asymptotisch}}{\sim} N(0, 1),$$

für große n , was bedeutet, dass

$$\mathbb{P}\left(\frac{\hat{\theta}_n - \theta}{\frac{\sigma}{\sqrt{n}}} \leq x\right) \rightarrow \Phi(x),$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist. Damit kann mit der Symmetrie der Normalverteilung gefolgert werden, dass das Intervall

$$\left(\hat{\theta}_n - a\frac{\sigma}{\sqrt{n}}, \hat{\theta}_n + b\frac{\sigma}{\sqrt{n}}\right), \text{ für } 0 \leq a, b < \infty,$$

μ mit Wahrscheinlichkeit $\Phi(b) - \Phi(-a)$ für $n \rightarrow \infty$ beinhaltet. Womit man a und b so wählen kann, dass diese Wahrscheinlichkeit ein bestimmtes Konfidenzlevel $1 - p$, für ein $p > 0$ hat. Die Länge des Intervalls wird minimiert, wenn man es symmetrisch um die Null wählt, also $a = b = z_{\frac{p}{2}}$, wobei $1 - \Phi(z_{\frac{p}{2}}) = \Phi(-z_{\frac{p}{2}}) = \frac{p}{2}$. Damit ist das Intervall

$$\hat{\theta}_n \pm z_{\frac{p}{2}} \frac{\sigma}{\sqrt{n}}$$

ein asymptotisches Konfidenzintervall für den Erwartungswert zum Konfidenzlevel $1 - p$.

Da σ meist nicht bekannt ist, wird es in der Praxis durch den konsistenten Schätzer der Standardabweichung

$$s_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2},$$

ersetzt. Da für eine Folge von Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$, welche in Verteilung gegen eine Zufallsvariable X konvergiert, und eine weitere Folge von Zufallsvariablen $(Y_n)_{n \in \mathbb{N}}$, welche in Verteilung gegen eine Konstante $c \in \mathbb{R}$ konvergiert, nach dem Lemma von Slutsky 2.3 gilt, dass $X_n + Y_n \rightarrow X + c$ und $X_n Y_n \rightarrow cX$ in Verteilung, konvergiert auch

$$\frac{\hat{\theta}_n - \theta}{\frac{s_n}{\sqrt{n}}} \rightarrow N(0, 1) \text{ in Verteilung.}$$

Daher ist das Konfidenzintervall

$$\hat{\theta}_n \pm z_{\frac{p}{2}} \frac{s_n}{\sqrt{n}}$$

auch asymptotisch gültig für das Konfidenzlevel $1 - p$.

Die Konvergenzrate der Monte Carlo Schätzung ist, wenn man das Konfidenzintervall betrachtet, der Ordnung $O(n^{-\frac{1}{2}})$. Diese Konvergenzrate ist

unabhängig von der Dimension. Die Konvergenz kann verbessert werden, wenn man die Varianz beziehungsweise geschätzte Varianz verkleinert. Im Kapitel 5.4 werden Methoden erläutert, welche die Varianz des Schätzers vermindert. Zuvor werden jedoch noch Methoden betrachtet, welche verwendet werden können um Zufallsvariablen und Zufallspfade zu simulieren.

5.2 Simulation von Zufallsvariablen

Mathematische Software Pakete bieten typischerweise Methoden zur Simulation von Zufallsvariablen aus bekannten Verteilungen an. Insbesondere die Möglichkeit eine Gleichverteilung $U(0, 1)$ zu simulieren, erlaubt es, daraus eine große Anzahl anderer Verteilungen zu berechnen. In Matlab kann man mit dem Befehl `rand(n,1)` und in R mit `runif(n, min=0, max=1)` eine Stichprobe von Zufallsvariablen $U \sim U(0, 1)$ der Größe n generieren. Die Stichproben werden, wie in [ABLM13] beschrieben, durch einen deterministischen Algorithmus, welcher eine Gleichverteilung gut imitiert, generiert. Diese Methoden nennt man „Pseudo-Random-Number-“Algorithmus, und die daraus resultierenden Zufallsvariablen nennt man „Pseudo“-Zufallszahlen.

Es werden nun einige Methoden vorgestellt, wie man eine Stichprobe gleichverteilter Zufallsvariablen zu einer Stichprobe aus anderen Verteilungen transformieren kann. Bei der Ausarbeitung dieser Methoden orientiert man sich größtenteils an [Gla04].

Inversionsmethode Angenommen man möchte eine Zufallsvariable X mit Verteilungsfunktion F_X generieren. Lemma 2.3 besagt, dass für eine gleichverteilte Zufallsvariable $U \sim U(0, 1)$ folgt, dass $F_X^{\leftarrow}(U)$ die gleiche Verteilung hat wie X , wobei F_X^{\leftarrow} die verallgemeinerte Inverse von F_X ist. Damit kann man mit dieser Methode, welche im folgenden Inversionsmethode genannt wird, eine Stichprobe für eine Zufallsvariable X aus einer Stichprobe von gleichverteilten Pseudo-Zufallszahlen bekommen.

Simulation von eindimensional normalverteilten Zufallsvariablen

Eine einfach zu implementierende Methode ist die Box-Muller Methode. Wenn $U_1, U_2 \sim U(0, 1)$ eine Stichprobe von zwei unabhängigen gleichverteilten Zufallsvariablen, dann gehört sowohl $Y_1 = \sqrt{-2 \log(U_1)} \sin(2\pi U_2)$ und $Y_2 = \sqrt{-2 \log(U_1)} \cos(2\pi U_2)$ zu einer $N(0, 1)$ -Verteilung.

Dass gilt, denn sind Y_1, Y_2 unabhängig standardnormalverteilt, dann ist $R = Y_1^2 + Y_2^2$ Chi-Quadrat verteilt mit zwei Freiheitsgraden. Über die charakteristische Funktion kann man leicht nachrechnen, dass das einer Exponentialverteilung mit Mittelwert 2 gleichzusetzen ist, was bedeutet

$$\mathbb{P}(R \leq x) = 1 - \exp\left(-\frac{x}{2}\right).$$

Hat man R gegeben, dann ist es möglich den Punkt (Y_1, Y_2) als einen gleichverteilten Punkt auf dem Kreis mit Radius \sqrt{R} und Mittelpunkt im Ur-

sprung zu betrachten. Um einen Zufallspunkt auf einem Kreis zu simulieren, kann man einen Zufallswinkel zwischen 0 und 2π generieren und diesen Zufallswinkel auf den Kreis projizieren.

1. Simuliere R über die Inversionsmethode, das heißt $R = -2 \log(U_1)$, wobei $U_1 \sim U(0, 1)$ gilt.
2. Simuliere einen Zufallswinkel durch $V = 2\pi U_2$ für $U_2 \sim U(0, 1)$ unabhängig von U_1 .
3. Der Punkt des Kreises hat dann die Koordinaten

$$\left(\sqrt{R} \cos(V), \sqrt{R} \sin(V) \right).$$

Damit sind die Zufallsvariablen $Y_1 = \sqrt{-2 \log(U_1) \sin(2\pi U_2)}$ und $Y_2 = \sqrt{-2 \log(U_1) \cos(2\pi U_2)}$ unabhängig standardnormalverteilte Zufallsvariablen.

In Matlab kann man eine Stichprobe standardnormalverteilter Zufallsvariablen der Größe n mit der Funktion `randn(n,1)` generieren.

Simulation von d-dimensionalen normalverteilten Zufallsvariablen

Sei $Z \sim N_d(\mu, \Sigma)$, wobei $\mu \in \mathbb{R}^d$ der Mittelwertvektor und $\Sigma \in \mathbb{R}^{d \times d}$ die Kovarianzmatrix ist. Für $X \sim N_d(0, I_d)$ mit I_d die d-dimensionale Einheitsmatrix, welche eine Stichprobe von d unabhängigen $N(0, 1)$ -verteilten Zufallsvariablen voraussetzt, gilt $Z \sim AX + \mu$, dabei wird A so festgelegt, dass $\Sigma = AA^T$ mit $A = (a_{ij})$ wird aus $\Sigma = (\sigma_{ij})$ durch die Choleski-Zerlegung berechnet:

$$\begin{cases} a_{11} = \sqrt{\sigma_{11}} \\ a_{ij} = 0, & \text{für } j > i \\ a_{ij} = \frac{1}{a_{jj}} (\sigma_{ij} - \sum_{k=1}^{j-1} a_{ik} a_{jk}), & \text{für } 1 \leq j < i \leq d \\ a_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{j-1} a_{ik}^2} \end{cases}$$

Die Simulation einer normalverteilten Stichprobe benötigt man zum Beispiel für die Simulation einer Brownschen Bewegung. Daher werden nun einige Eigenschaften der mehrdimensionalen Normalverteilung angeführt, welche später dann benötigt werden.

- Lineare Transformationseigenschaft:

Die lineare Transformation eines normalverteilten Zufallsvektor ist wieder normalverteilt:

$$X \sim N(\mu, \Sigma) \implies AX \sim N(A\mu, A\Sigma A^T), \quad (21)$$

für einen Vektor $\mu \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d \times d}$ und eine Matrix $A \in \mathbb{R}^{k \times d}$ für ein $k \in \mathbb{N}$ beliebig.

- Bedingte Verteilung normalverteilter Zufallsvektoren:
Sei $(X_{[1]}, X_{[2]})$ ein partitionierter Vektor (wobei $X_{[i]}$ jeweils auch ein Vektor sein kann), welcher multivariat normalverteilt ist mit

$$\begin{pmatrix} X_{[1]} \\ X_{[2]} \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_{[1]} \\ \mu_{[2]} \end{pmatrix}, \begin{pmatrix} \Sigma_{[11]} & \Sigma_{[12]} \\ \Sigma_{[21]} & \Sigma_{[22]} \end{pmatrix} \right),$$

wobei $\Sigma_{[22]}$ vollen Rang hat. Die Dimension von $\mu_{[i]}$ und $\Sigma_{[ij]}$ stimmen für $i, j = 1, 2$ mit jenen von $X_{[i]}$ überein. Dann ist die Verteilung des Vektors $X_{[1]}$ bedingt auf $X_{[2]} = x$ gleich

$$(X_{[1]} | X_{[2]} = x) \sim N(\mu_{[21]}, \Sigma_{[221]}) \quad (22)$$

mit

$$\mu_{[21]} = \mu_{[1]} + \Sigma_{[12]}(\Sigma_{[22]})^{-1}(x - \mu_{[2]}) \quad (23)$$

und

$$\Sigma_{[11]} - \Sigma_{[12]}(\Sigma_{[22]})^{-1}\Sigma_{[21]}. \quad (24)$$

- Momentenerzeugende Funktion:
Wenn $X \sim N(\mu, \Sigma)$, wobei X d -dimensional ist, dann ist die momentenerzeugende Funktion wie folgt gegeben:

$$E[\exp(\theta^T X)] = \exp\left(\mu^T \theta + \frac{1}{2} \theta^T \Sigma \theta\right) \quad (25)$$

für alle $\theta \in \mathbb{R}^d$.

Für eine detaillierte Beschreibung der multivariaten Normalverteilung und deren Eigenschaften sowie den Beweisen siehe [FHT96].

5.3 Simulation von Stichprobenpfade

5.3.1 Brownsche Bewegung eindimensional

Um einen Zufallspfad einer Brownschen Bewegung nach Definition 2.13 zu generieren, sei $0 = t_0 < t_1 < \dots < t_n = T$ eine fixe Partition von $[0, T]$. Da die Zuwächse der Brownschen Bewegung unabhängig und normalverteilt sind, ist es naheliegend W über ihre Zuwächse zu simulieren. Dazu simuliere man Z_1, \dots, Z_n unabhängige standardnormalverteilte Zufallsvariablen. Für eine Brownsche Bewegung setze $W_{t_0} = 0$. Nachfolgende Werte, können dann wie folgt generiert werden.

$$W_{t_{i+1}} = W_{t_i} + \sqrt{t_{i+1} - t_i} Z_{t_{i+1}}, \text{ für } i = 0, \dots, n-1. \quad (26)$$

Eine andere Methode zur Simulation einer Brownschen Bewegung verwendet den Ansatz, dass der Vektor $(W_{t_1}, \dots, W_{t_n})^T$ eine lineare Transformation des Vektors $(W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}})^T$ ist, wobei A ist

eine untere Dreiecksmatrix mit $(a_{ij}) = 1$ für $1 \leq j \leq i \leq n$. Da die Zuwächse unabhängig und normalverteilt sind, und eine lineare Transformation eines normalverteilten Zufallsvektor wieder normalverteilt ist, folgt, dass W_{t_1}, \dots, W_{t_n} eine multivariate Normalverteilung hat. Also kann man die Simulation der Brownschen Bewegung, wie einen speziellen Fall zur Erzeugung eines mehrdimensionalen normalverteilten Zufallsvektor behandeln. Dazu benötigt man den Erwartungsvektor und die Kovarianzmatrix. Es gilt $E[W_{t_i}] = 0$ für alle $i = 1, \dots, n$. Für die Kovarianzmatrix sei $0 \leq s < t \leq T$. Verwendet man die Unabhängigkeit der Zuwächse folgt:

$$\begin{aligned} \text{Cov}[W_s, W_t] &= \text{Cov}[W_s, W_s + (W_t - W_s)] \\ &= \text{Cov}[W_s, W_s] + \text{Cov}[W_s, W_t - W_s] = s + 0 = s. \end{aligned}$$

Für die Kovarianzmatrix C von $(W_{t_1}, \dots, W_{t_n})$ gilt daher $C_{ij} = \min(t_i, t_j)$ für $1 \leq i, j \leq n$.

Brownsche Brücke Eine dritte Methode für die Simulation der Brownschen Bewegung ist die Brownsche Brücke. Bei dieser Methode simuliert man die Brownsche Bewegung nicht von links nach rechts, sondern es wird die bedingte Verteilung der einzelnen Punkte betrachtet, wenn man außer dem Startwert einen weiteren Wert gegeben hat. Damit kann man, die einzelnen Werte in einer beliebigen Reihenfolge berechnen.

Sei zuerst $0 < u < s < t$. Es wird das Problem betrachtet W_s unter der Bedingung $W_u = x$ und $W_t = y$ zu berechnen. Dazu verwendet man die Formel (22) für die bedingte Verteilung eines normalverteilten Zufallsvektors. Die Verteilung des Zufallsvektors ist wie folgt gegeben

$$\begin{pmatrix} W_u \\ W_s \\ W_t \end{pmatrix} \sim N \left(0, \begin{pmatrix} u & u & u \\ u & s & s \\ u & s & t \end{pmatrix} \right).$$

Um die bedingte Verteilung zu berechnen, vertauscht man die Vektoreinträge wie folgt:

$$\begin{pmatrix} W_s \\ W_u \\ W_t \end{pmatrix} \sim N \left(0, \begin{pmatrix} s & u & s \\ u & u & u \\ s & u & t \end{pmatrix} \right).$$

Nun kann der Erwartungswert und die Varianz mit den Formeln (23) und (24) berechnet werden. Es gilt

$$E[W_s | W_u = x, W_t = y] = \frac{(t-s)x + (s-u)y}{t-u}$$

und

$$\text{Var}[W_s | W_u = x, W_t = y] = \frac{(s-u)(t-s)}{t-u}.$$

Betrachtet man allgemeiner, dass $W_{s_1} = x_1, \dots, W_{s_k} = x_k$ für $s_1 < \dots < s_k$ des Brownschen Pfades bereits berechnet wurden. Und man möchte W_s für $s_i < s < s_{i+1}$ berechnen. Die Markoveigenschaft der Brownschen Bewegung besagt, dass hat man W_{s_i} gegeben, dann ist W_s unabhängig von allen W_t mit $t < s_i$ und genauso gilt, dass wenn man $W_{s_{i+1}}$ gegeben hat, dann ist W_s unabhängig von allen W_t mit $t > s_{i+1}$, was eine Konsequenz der Unabhängigkeit der Zuwächse ist. Damit gilt

$$(W_s \mid W_{s_1} = x_1, \dots, W_{s_k} = x_k) = (W_s \mid W_{s_i} = x_i, W_{s_{i+1}} = x_{i+1}).$$

Für die bedingte Verteilung von W_s gilt dann

$$\begin{aligned} (W_s \mid W_{s_1} = x_1, \dots, W_{s_k} = x_k) &\sim \\ &\sim N\left(\frac{(s_{i+1} - s)x_i + (s - s_i)x_{i+1}}{s_{i+1} - s_i}, \frac{(s_{i+1} - s)(s - s_i)}{s_{i+1} - s_i}\right) \end{aligned}$$

und daher

$$W_s = \frac{(s_{i+1} - s)x_i + (s - s_i)x_{i+1}}{s_{i+1} - s_i} + \sqrt{\frac{(s_{i+1} - s)(s - s_i)}{s_{i+1} - s_i}} Z,$$

wobei $Z \sim N(0, 1)$ unabhängig von allen W_{s_1}, \dots, W_{s_k} . Durch wiederholte Anwendung dieser Formel, erhält man eine Stichprobe der Komponenten des Vektors $(W_{t_1}, \dots, W_{t_n})$ in einer beliebigen Reihenfolge.

Diese alternativen Methoden kann man auch auf die Brownsche Bewegung mit Drift und die geometrische Brownsche Bewegung übertragen. Darauf wird aber im Folgenden nicht näher eingegangen.

5.3.2 Brownsche Bewegung mit Drift

Hat man eine fixe Menge von Punkten $0 < t_1 < \dots < t_n$ gegeben, kann man eine Brownsche Bewegung mit Drift, wie in Definition 5 für ein gegebenes $X_0 \in \mathbb{R}$ durch

$$X_{t_{i+1}} = X_{t_i} + \mu(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i}Z_{i+1}, \text{ für } i = 0, \dots, n - 1 \quad (27)$$

simuliert werden, wobei die Z_1, \dots, Z_n eine unabhängige Stichprobe aus einer Standardnormalverteilung ist.

5.3.3 Geometrische Brownsche Bewegung

Es wird nun ein einfacher Algorithmus für die Simulation einer geometrischen Brownschen Bewegung hergeleitet, welche in (6) definiert ist. Die Tatsache, dass die Zuwächse der Brownschen Bewegung unabhängig und normalverteilt sind, vereinfacht die Simulation der Werte von S . Es sei

$0 = t_0 < t_1 < \dots < t_n$, dann kann man eine geometrische Brownsche Bewegung durch

$$S_{t_{i+1}} = S_{t_i} \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) (t_{i+1} - t_i) + \sigma \sqrt{t_{i+1} - t_i} Z_{i+1} \right) \quad (28)$$

für $i=0, \dots, n-1$, wobei Z_1, \dots, Z_n unabhängig standardnormalverteilte Zufallsvariablen sind. Das ist äquivalent dazu auf beiden Seiten von (27) die Exponentialfunktion anzuwenden und μ wird durch $\mu - \frac{1}{2} \sigma^2$ ersetzt.

5.3.4 Mehrdimensionale Brownsche Bewegung

Für die Simulation einer d-dimensionalen Brownschen Bewegung und Brownschen Bewegung mit Drift wie sie in 2.16 und 2.17 definiert wurden, generiere man zuerst Z_1, Z_2, \dots, Z_n unabhängige $N(0, I)$ -verteilte d-dimensionale Zufallsvektoren. Für eine fixe Menge von Punkten $0 = t_0 < t_1 < \dots < t_n$ setze $W_0 = 0$ und

$$W_{t_{i+1}} = W_{t_i} + \sqrt{t_{i+1} - t_i} Z_{i+1} \text{ für } i = 0, \dots, n-1. \quad (29)$$

Das ist äquivalent dazu einen eindimensionalen Zufallspfad für jede Koordinate von W zu konstruieren.

Um eine d-dimensionale Brownsche Bewegung mit Drift zu simulieren, berechnet man zuerst A , sodass $AA^T = \Sigma$, mit der Choleski-Zerlegung. Ist $A \in \mathbb{R}^{d \times k}$ seien Z_1, Z_2, \dots, Z_n standardnormalverteilte k-dimensionale Zufallsvektoren. Setze $X_0 = x \in \mathbb{R}^d$ und

$$X_{t_{i+1}} = X_{t_i} + \mu(t_{i+1} - t_i) + \sqrt{t_{i+1} - t_i} A Z_{i+1} \text{ für } i = 0, \dots, n-1. \quad (30)$$

Für die Simulation einer d-dimensionalen geometrischen Brownschen Bewegung beachte man zuerst, dass nach Definition 7

$$S_t^i = S_0^i \exp \left(\left(\mu_i - \frac{1}{2} \sigma^2 \right) t + \sigma_i W_t^i \right) \text{ für } i = 1, \dots, d.$$

gilt. Definiert man eine $(d \times d)$ -Matrix Σ durch $\Sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$, dann ist $(\sigma_1 W^1, \dots, \sigma_d W^d)$ eine d-dimensionale Brownsche Bewegung mit Drift, wobei die Kovarianzmatrix gleich Σ ist. Ein solcher Prozess kann durch AB_t geschrieben werden, wobei B eine d-dimensionale Brownsche Bewegung und A eine Matrix ist, welche $AA^T = \Sigma$ erfüllt. Damit kann man (7) als

$$\frac{dS_t^i}{S_t^i} = \mu_i dt + \sum_{j=1}^d A_{ij} dW_t^j \text{ für } i = 1, \dots, d$$

schreiben. Also kann man eine d-dimensionalen geometrischen Brownschen Bewegung mit Driftvektor (μ_1, \dots, μ_d) und Volatilitätsmatrix Σ durch

$$S_{t_{k+1}}^i = S_{t_k}^i \exp\left(\left(\mu_i - \frac{1}{2}\sigma_i^2\right)(t_{k+1} - t_k) + \sqrt{t_{k+1} - t_k} \sum_{j=1}^d A_{ij} Z_{k+1}^j\right) \quad (31)$$

für $i = 1, \dots, d$ und $k = 0, \dots, n-1$, simulieren, wobei $Z_k = (Z_k^1, \dots, Z_k^d) \sim N(0, I)$ und Z_1, \dots, Z_d unabhängig sind.

Auch im d-dimensionalen Fall ist der Algorithmus für die geometrische Brownsche Bewegung äquivalent zur Anwendung der Exponentialfunktion auf beiden Seiten des Algorithmus für die Brownsche Bewegung mit Drift, wobei μ_i durch $\mu_i - \frac{1}{2}\sigma_i^2$ ersetzt wird.

5.4 Methoden zur Varianzreduktion

Es gilt, dass der Approximationsfehler der Monte Carlo Methode in Wahrscheinlichkeit der Ordnung $O(n^{-\frac{1}{2}})$ ist, daher benötigt man für die Verbesserung der Genauigkeit der Schätzung um eine Dezimalstelle ein Anwachsen der Anzahl der Simulationen n um einen Faktor 100.

Man könnte die Konvergenzrate verbessern, indem man die Varianz des Monte Carlo Schätzers reduziert. Das ist möglich, wenn man z.B. eine Zufallsvariable Z' findet mit $E[g(Z')] = E[g(Z)] = \alpha$ und $Var[g(Z')] \leq Var[g(Z)]$, dann kann eine Schätzung von $E[g(Z)]$ durch Simulation von Z' erreicht werden. Das Problem ist ein Z' zu finden, welches diese Forderung erfüllt. Es ist nur dann sinnvoll einen Algorithmus zu implementieren, welcher die Varianz reduziert, wenn die $Var[g(Z')]$ signifikant kleiner ist als $Var[g(Z)]$. Ist die Reduktion der Varianz nicht in einem ausreichenden Ausmaß gegeben, wäre es vielleicht effizienter die Anzahl der Simulationen um diesen Grad zu erhöhen. Es werden nun einige Methoden der Varianzreduktion vorgestellt. Großteils orientiert sich dieses Kapitels an [Gla04], wobei zum Vergleich auch [ABLM13] für die Ausarbeitung studiert wurde. Von diesen Methoden werden, dann jene, deren Anwendung sinnvoll ist in den Beispielen in Kapitel 9 verglichen.

5.4.1 Antithetische Zufallsvariable

Bei dieser Methode versucht man die Varianz zu reduzieren, indem man eine negative Abhängigkeit bei Stichprobenpaaren einführt. Die gängigste Form dieser Methode basiert darauf, dass für eine gleichverteilte Zufallsvariable $U \sim U(0, 1)$ auch $1-U \sim U(0, 1)$ gilt. Generiert man eine Stichprobe, welche als Eingabe auf $(0, 1)$ gleichverteilte U_1, \dots, U_n verwendet, kann eine zweite Stichprobe unter den gleichen Voraussetzungen generiert werden, welche $1 - U_1, \dots, 1 - U_n$ als Eingabe verwendet. Die Variablen U_i und $1 - U_i$, für $i = 1, \dots, n$ formen ein antithetisches Paar in dem Sinne, dass ein großer

Wert des einen durch einen kleinen Wert des anderen relativiert wird. Wenn man in der Simulation der ersten Stichprobe eine Ausgabe hat, welcher sehr groß oder sehr klein ist, wird dieser durch den Wert der Ausgabe der zweiten Stichprobe ausgeglichen. Indem man diese antithetischen Paare verwendet, möchte man die Varianz reduzieren.

Diese Beobachtung kann man durch die Inversionsmethode auf andere Verteilungen erweitern, da $F^{\leftarrow}(U)$ und $F^{\leftarrow}(1-U)$ die selbe Verteilung F haben, aber antithetisch zueinander sind, da F^{\leftarrow} monoton ist. Für eine Verteilung, welche symmetrisch um 0 ist, haben $F^{\leftarrow}(U)$ und $F^{\leftarrow}(1-U)$ die selbe Größe aber andere Vorzeichen.

Insbesondere können bei einer Simulation, welche von unabhängigen standardnormalverteilten Zufallsvariablen abhängt, die antithetischen Paare $Z_i \sim N(0, 1)$ und $-Z_i \sim N(0, 1)$ für $i = 1, \dots, n$ verwendet werden. Dabei hängt es nicht davon ab ob die normalverteilten Zufallsvariablen mit der Inversionsmethode oder durch einen anderen Algorithmus generiert werden.

Das Ziel ist es nun wieder den $E[g(Z)]$ zu schätzen, dazu setzt man wieder für die Vereinfachung der Notation $Y = g(Z)$. Es wird zuerst eine Stichprobe von antithetischen Paaren $(Y_1, \tilde{Y}_1), \dots, (Y_n, \tilde{Y}_n)$ generiert, welche folgende Eigenschaften erfüllen:

- Die Paare $(Y_1, \tilde{Y}_1), \dots, (Y_n, \tilde{Y}_n)$ sind i.i.d.
- Für jedes i haben Y_i und \tilde{Y}_i die selbe Verteilung, sind aber normalerweise nicht unabhängig.

Der antithetische Variablen Schätzer ist dann der Durchschnitt aller $2n$ Beobachtungen.

$$\hat{\alpha}_Y^{AV} = \frac{1}{2n} \left(\sum_{i=1}^n Y_i + \sum_{i=1}^n \tilde{Y}_i \right) = \frac{1}{n} \sum_{i=1}^n \left(\frac{Y_i + \tilde{Y}_i}{2} \right)$$

Die Darstellung ganz rechts zeigt, dass der Schätzer $\hat{\alpha}_Y^{AV}$ das Stichprobenmittel der n unabhängigen Beobachtungen

$$\left(\frac{Y_1 + \tilde{Y}_1}{2} \right), \dots, \left(\frac{Y_n + \tilde{Y}_n}{2} \right)$$

ist.

Wenn man davon ausgeht, dass die Simulation der n antithetischen Paaren gleich aufwendig ist, wie die Simulation von $2n$ unabhängigen Stichprobenpunkten, dann würde sich diese Methode erst bezahlt machen, wenn

$$Var[\hat{\alpha}_Y^{AV}] < Var\left[\frac{1}{2n} \sum_{i=1}^{2n} Y_i\right]$$

$$\begin{aligned} &\Leftrightarrow \frac{1}{4n} \text{Var} [Y_i + \tilde{Y}_i] < \frac{1}{2n} \text{Var} [Y_i] \\ &\Leftrightarrow \text{Var} [Y_i + \tilde{Y}_i] < 2\text{Var} [Y_i]. \end{aligned}$$

Sieht man sich die linke Seite dieser Ungleichung genauer an gilt

$$\begin{aligned} \text{Var} [Y_i + \tilde{Y}_i] &= \text{Var} [Y_i] + \text{Var} [\tilde{Y}_i] + 2\text{Cov} [Y_i, \tilde{Y}_i] \\ &= 2\text{Var} [Y_i] + 2\text{Cov} [Y_i, \tilde{Y}_i], \end{aligned}$$

da Y_i und \tilde{Y}_i für $i = 1, \dots, n$ gleich verteilt sind. Daher erreicht man eine Reduktion der Varianz, wenn

$$\text{Cov} [Y_i, \tilde{Y}_i] < 0,$$

also muss aus einer negativen Abhängigkeit der Eingabe, eine negative Korrelation der Ausgabe gefolgert werden können. Eine hinreichende Bedingung dafür ist die Monotonie der Abbildung, welche die Eingabe auf die Ausgabe abbildet.

Um dies etwas anschaulicher wiederzugeben, sei X eine Inputvariable. Angenommen $Y = f(X)$ ist eine monoton wachsende Funktion und $\tilde{Y} = g(X)$ ist eine monoton fallende Funktion, dann gilt $\text{Cov} [Y, \tilde{Y}] \leq 0$, denn ist X^* eine von X unabhängige und identisch verteilte Zufallsvariable, dann gilt fast sicher

$$\begin{aligned} 0 &\geq E [(f(X) - f(X^*))(g(X) - g(X^*))] = \\ &E [f(X)g(X)] - E [f(X^*)g(X)] - E [f(X)g(X^*)] + E [f(X^*)g(X^*)] \\ &\stackrel{X, X^* \text{ u.a. und identisch verteilt}}{=} 2E [f(X)g(X)] - 2E [f(X)] E [g(X)] \\ &= 2E [Y\tilde{Y}] - 2E [Y] E [\tilde{Y}] = 2\text{Cov} [Y, \tilde{Y}]. \end{aligned}$$

Betrachtet man eine Stichprobe gleichverteilte Zufallsvariablen U_1, \dots, U_n bzw. normalverteilte Zufallsvariablen Z_1, \dots, Z_n , und ist $Y = f(U_1, \dots, U_n)$ bzw. $Y = f(Z_1, \dots, Z_n)$ eine monoton wachsende Funktion, dann ist $\tilde{Y} = f(1 - U_1, \dots, 1 - U_n)$ bzw. $\tilde{Y} = f(-Z_1, \dots, -Z_n)$ eine monoton fallende Funktion von (U_1, \dots, U_n) bzw. (Z_1, \dots, Z_n) .

Die antithetischen Paare der Normalverteilung und der Gleichverteilung teilen eine weitere wichtige Eigenschaft. Nämlich ist der Durchschnitt eines antithetischen Paares genau der Erwartungswert. Also wenn der Output eine lineare Funktion von gleichverteilten bzw. normalverteilten Zufallsvariablen ist, dann verschwindet die Varianz komplett, was an der Symmetrie dieser beiden Verteilungen liegt.

Zerlegung der Varianz Antithetische Variablen eliminieren den Anteil der Varianz, welcher durch den antisymmetrischen Part des Integranden entsteht. Diese Behauptung wird im Falle von standardnormalverteilten Zufallsvariablen erarbeitet. Durch eine leichte Modifikation der folgenden Überlegungen lassen sich diese auch auf andere symmetrische Verteilungen übertragen.

Sei $Y = f(Z)$ wobei $Z = (Z_1, \dots, Z_d) \sim N(0, 1)$. Zerlegt man die Funktion in ihren symmetrischen und antisymmetrischen Teil, und setze

$$f_1(z) = \frac{f(z) + f(-z)}{2}, \quad f_2(z) = \frac{f(z) - f(-z)}{2}.$$

Das ist eine orthogonale Zerlegung von f , in diesem Sinne das f_1 und f_2 unkorreliert sind, denn da Z und $-Z$ identisch verteilt sind gilt

$$\begin{aligned} E[f_1(Z)f_2(Z)] &= \frac{1}{4}E[f^2(Z) + f(Z)f(-Z) - f(Z)f(-Z) - f^2(-Z)] \\ &= \frac{1}{4}(E[f^2(Z)] - E[f^2(-Z)]) = 0 \end{aligned}$$

und gleich folgt auch

$$E[f_1(Z)]E[f_2(Z)] = 0.$$

Daher kann die Varianz in

$$\text{Var}[f(Z)] = \text{Var}[f_1(Z)] + \text{Var}(f_2(Z))$$

aufgeteilt werden. Die $\text{Var}[f_1(Z)]$ ist genau die Varianz des Schätzers, welchen man erhält, wenn man das antithetische Paar $(Z, -Z)$ für die Schätzung verwendet. Damit kann man folgern, dass die Varianz einer antisymmetrischen Funktion ($f = f_2$) durch die Anwendung von antithetischen Paaren zur Gänze verschwindet, wobei die Varianz sich nicht verkleinert, wenn man eine symmetrische Funktion ($f = f_1$) hat.

Im folgenden Beispiel wird diese Methode auf die Schätzung eines Integrals angewendet.

Beispiel 5.1. *Es wird versucht das Integral*

$$\int_0^1 \frac{1}{1+x} dx$$

zu schätzen, was auch als Erwartungswert $E[f(U)]$ für U eine standardgleichverteilte Zufallsvariable und $f(x) = \frac{1}{1+x}$ betrachtet werden kann. Der exakte Wert dieses Integrals ist

$$\ln(2) = 0.69315.$$

Mittels Monte Carlo Simulation in Matlab erhält man für $n = 2000$ folgende Werte

$n = 2000$	Schätzung	Varianz
klassische Monte Carlo Schätzung	0.69068	0.01969
Schätzung mit antithetischen Zufallszahlen	0.069327	0.0006055

Tabelle 1: Integralschätzung klassisch und antithetische Stichprobe

Man kann sehen, dass die Varianz des antithetischen Schätzers signifikant kleiner ist und die Schätzung um vieles genauer als bei der klassischen Methode.

5.4.2 Kontrollierende Zufallsvariable

Diese Methode ist die am Häufigsten in der Praxis angewendete Methode. Man verwendet die Information über den Fehler einer Schätzung für einen bekannten Parameter, um den Fehler der Schätzung eines unbekanntes Parameters zu reduzieren. Ziel ist es wieder $E[g(Z)]$ zu schätzen, sei dazu o.B.d.A $g(Z) = Z$. Weiter sei Y eine Zufallsvariable, deren Mittel $E[Y]$ bekannt ist, und es gilt $Cov[Y, Z] \neq 0$. Man kann (Z, Y) gemeinsam simulieren und es gilt

$$E[Z] = E[Z - aY] + aE[Y]$$

für eine Konstante $a \neq 0$. Nun schätzen wir $E[Z - aY]$ anstelle von $E[Z]$. Dazu generiere n unabhängige Stichproben $(Z_1, Y_1), \dots, (Z_n, Y_n)$ des Vektors (Z, Y) , und berechne

$$Z_j(a) = Z_j - a(Y_j - E[Y]), \quad j = 1, \dots, n$$

und den neuen Schätzer

$$\hat{\alpha}_Z(a) := \frac{1}{n} \sum_{j=1}^n Z_j(a) = \hat{\alpha}_Z - a(\hat{\alpha}_Y - E[Y]),$$

welcher aus dem klassischen Monte Carlo Schätzer $\hat{\alpha}_Z = \frac{1}{n} \sum_{j=1}^n Z_j$ und dem beobachtbaren Schätzfehler $\hat{\alpha}_Y - E[Y] = \frac{1}{n} \sum_{j=1}^n Y_j - E[Y]$ zusammengesetzt ist. Da von unverzerrten Schätzern ausgegangen wird, gilt $E[\hat{\alpha}_Z(a)] = E[Z]$.

Seien nun σ_Y^2 und σ_Z^2 die Varianzen von Y und Z , und der Korrelationskoeffizient wird mit $\rho_{Y,Z}$ bezeichnet, wobei für den Korrelationskoeffizienten gilt

$$\rho_{Y,Z} = \frac{Cov[Y, Z]}{\sigma_Y \sigma_Z}$$

. Betrachtet man nun die Varianz des Schätzers ergibt diese:

$$Var[Z_j(a)] = \sigma_Z^2 - 2a\sigma_Y\sigma_Z\rho_{Y,Z} + a^2\sigma_Y^2 = nVar[\hat{\alpha}_Z(a)] \quad (32)$$

Da $Var[\hat{\alpha}_Z] = Var[\hat{\alpha}_Z(0)] = \frac{\sigma_Z^2}{n}$ hat der neue Schätzer $\hat{\alpha}_Z(a)$ kleinere Varianz als $\hat{\alpha}_Z$, wenn $a^2\sigma_Y < 2a\sigma_Z\rho_{Y,Z}$. Durch (32) ist es möglich den Wert von a , durch ableiten und Null setzen, so zu bestimmen, dass die Varianz des neuen Schätzers minimiert wird. Minimale Varianz erhält man für den Koeffizienten

$$a^* = \rho_{Y,Z} \frac{\sigma_Z}{\sigma_Y} = \frac{Cov[Y, Z]}{Var[Z]}.$$

Die Verbesserung der Varianz bei optimaler Wahl des Parameters ist durch

$$\frac{Var[\hat{\alpha}_Z(a^*)]}{Var[\hat{\alpha}_Z]} = 1 - 2a^* \frac{\sigma_Y}{\sigma_Z} \rho_{Y,Z} + (a^*)^2 \frac{\sigma_Y^2}{\sigma_Z^2} = 1 - \rho_{Y,Z}^2$$

gegeben. Man kann also sehen, dass eine stärkere Korrelation zwischen Z und Y zu einer kleineren Varianz führt.

Im Allgemeinen sind in der Praxis $Var[Z]$ und $\rho_{Y,Z}$ nicht bekannt. Man kann jedoch a^* durch den Schätzer

$$\hat{a}^* = \frac{\sum_{j=1}^n (y_j - \hat{\alpha}_Y)(z_j - \hat{\alpha}_Z)}{\sum_{j=1}^n (y_j - \hat{\alpha}_Y)^2}$$

ersetzen. Meistens erhält man mit dieser Methode eine signifikante Reduktion der Varianz und daher auch eine Verbesserung der Konvergenz.

Multiple Kontrolle Es wurde bis jetzt nur der Fall betrachtet, bei dem eine kontrollierende Variable verwendet wurde. Es ist jedoch auch möglich mehrere Variablen zur Kontrolle heranzuziehen. Dies erfordert eine Anpassung der Formeln, welche aber recht schnell hergeleitet ist.

Seien Z_i und $Y_i = (Y_i^{(1)}, \dots, Y_i^{(d)})$ für $i = 1, \dots, n$ eine generierte Stichprobe, wobei der Erwartungsvektor von Y bekannt ist. Man nehme nun an, dass die Vektoren (Y_i, Z_i) für $i = 1, \dots, n$ unabhängig und identisch verteilt sind mit Kovarianzmatrix

$$\begin{pmatrix} \Sigma_Y & \Sigma_{YZ} \\ \Sigma_{YZ}^T & \sigma_Z^2 \end{pmatrix}, \quad (33)$$

wobei $\Sigma_Y \in \mathbb{R}^{d \times d}$, $\Sigma_{YZ} \in \mathbb{R}^d$ und $\sigma_Z^2 > 0$ ist die Varianz von Z . Es wird vorausgesetzt, dass Σ_Y nicht singulär ist, da einige $Y^{(k)}$ sonst als Linearkombinationen der anderen dargestellt werden könnten. Wäre dies der Fall sollte man jene aus der Menge der kontrollierenden Zufallsvariablen entfernen.

Sei nun $\hat{\alpha}_Y$ der Vektor der Stichprobenmittelwerte der Kontrollen. Für fixes $b \in \mathbb{R}^d$ ist der neue Schätzer $\hat{\alpha}_Z(a)$ gegeben durch:

$$\hat{\alpha}_Z(a) = \hat{\alpha}_Z - a^T(\hat{\alpha}_Y - E[Y]).$$

Die Varianz ist, dann gleich wie im einfachen Fall durch

$$\text{Var} [Z_i - a^T(Y_i - E[Y])] = \sigma_Z - 2a^T \Sigma_{YZ} + b^T \Sigma_Y a$$

gegeben, welche an der Stelle

$$a^* = \Sigma_Y^{-1} \Sigma_{YZ}$$

minimiert wird.

In der Praxis muss man die Varianz und Korrelation durch ihre Standardschätzer ersetzen da sie nicht bekannt sind. In diesem Fall gilt

$$\hat{a}_n = S_Y^{-1} S_{YZ}$$

wobei $S_Y \in \mathbb{R}^{d \times d}$ mit

$$(S_Y)_{i,j} = \frac{1}{n-1} \left(\sum_{k=1}^n Y_k^{(j)} Y_k^{(i)} - N \hat{\alpha}_Y^{(j)} \hat{\alpha}_Y^{(i)} \right)$$

und $S_{YZ} \in \mathbb{R}^d$ mit

$$(S_{YZ})_i = \frac{1}{n-1} \left(\sum_{k=1}^n Y_k^{(i)} Z_k - N \hat{\alpha}_Y^{(i)} \hat{\alpha}_Z \right).$$

Man sollte bei der Wahl der Kontrollen bedenken, dass Kontrollen welche stark miteinander korrelieren S_Y fast singulär werden lassen. Bei der Wahl sollte man also immer bedacht vorgehen.

Zerlegung der Varianz Um die Methode effektiv anzuwenden, wird nun überlegt, welcher Teil der Varianz reduziert wird. Sei dazu (Z, Y) ein Zufallsvektor, wobei Y d -dimensional und Z eindimensional ist. Für ein beliebiges $a \in \mathbb{R}^d$ wähle ϵ so dass

$$Z = E[Z] + b^T(Y - E[Y]) + \epsilon$$

gilt. Wenn $a = a^* = \Sigma_Y^{-1} \Sigma_{YZ}$ gewählt wird, dann gilt

$$\text{Cov}[Y, \epsilon] = \text{Cov}[Y, Z] - (a^*)^T \text{Cov}[Y, Y] = 0.$$

Damit hat man

$$\text{Var}[Z] = \text{Var}[(a^*)^T Y] + \text{Var}[\epsilon] = \text{Var}[(a^*)^T Y] + \text{Var}[Z - (a^*)^T Y].$$

Daraus kann abgeleitet werden, dass wenn Y als Kontrolle verwendet wird bleibt $\text{Var}[\epsilon]$. Wäre es möglich Z als Linearkombination der Variablen Y zu schreiben, würde man die gesamte Varianz eliminieren können, jedoch wäre es dann auch möglich $E[Z]$ direkt als eine Linearkombination der bekannten Erwartungswerte von $E[Y^{(i)}]$ für $i = 1, \dots, d$ zu berechnen. Es ist also wünschenswert die Kontrolle so zu wählen, dass sich Z so gut wie möglich in Termen von Y darstellen lässt.

5.4.3 Importance Sampling

Die Idee dieser Methode ist es zu einem Wahrscheinlichkeitsmaß zu wechseln, welches wichtige Stichprobenpunkte (oder Pfade) öfter generiert. Was eine signifikante Verbesserung bei der Schätzung von seltenen Ereignissen bringen kann. Es wird hier nur der Fall betrachtet, dass Z eine Wahrscheinlichkeitsdichte $f : \mathbb{R}^s \rightarrow \mathbb{R}^+$ hat. Es gilt, dass

$$\alpha = E[g(Z)] = \int_{\mathbb{R}^s} g(z)f(z)dz$$

Sei $f_I : \mathbb{R}^s \rightarrow \mathbb{R}^+$ eine zweite Wahrscheinlichkeitsdichte mit der Eigenschaft, dass wenn $f(z) > 0$ gilt auch $f_I(z) > 0$ für alle $z \in \mathbb{R}^s \setminus \{z \mid g(z) = 0\}$. Dann kann man

$$\alpha = \int_{\mathbb{R}^s} g(z) \frac{f(z)}{f_I(z)} f_I(z) dz$$

schreiben, was als Erwartung in Bezug auf die Dichte f_I verstanden werden kann, d.h.

$$\alpha = E_I \left[g(z) \frac{f(z)}{f_I(z)} \right],$$

wobei der Index I bei dem Erwartungswert anzeigt, dass Z nun bzgl. der Dichte f_I betrachtet wird.

$\frac{f(z)}{f_I(z)}$ wird Radon-Nikodym Dichte oder Likelihood Ratio genannt.

Man kann nun Monte Carlo Simulation verwenden um N Stichprobenpunkte Z_1, \dots, Z_n von f_I zu generieren, um daraus den unverzerrten Schätzer

$$\hat{\alpha}_I = \frac{1}{n} \sum_{i=1}^n g(Z_i) \frac{f(Z_i)}{f_I(Z_i)} \tag{34}$$

zu berechnen. Diese Methode bringt eine Reduktion der Varianz, wenn f_I so gewählt wird, dass

$$E_I \left[\left(g(Z) \frac{f(Z)}{f_I(Z)} \right)^2 \right] \stackrel{\text{Maßwechsel}}{=} E \left[(g(Z))^2 \frac{f(Z)}{f_I(Z)} \right] < E \left[(g(Z))^2 \right].$$

Es ist von der Wahl der Dichte abhängig, ob die Varianz größer oder kleiner wird. Es ist sogar möglich, dass die Varianz unendlich groß wird. Nimmt man an, das g nicht negativ ist, dann ist das Produkt $g(Z)f(Z)$ auch nicht negativ. Damit wäre es möglich das Produkt zu normalisieren, dass es eine Wahrscheinlichkeitsdichte ist, dann könnte man f_I gleich dieser Wahrscheinlichkeitsdichte wählen. f_I ist dann proportional zu gf , dadurch erreicht man einen Schätzer $\hat{\alpha}_I$ mit Varianz 0. Jedoch wird α benötigt um gf zu normalisieren, und da α aber die Größe ist, welche geschätzt werden soll, ist dieser Ansatz in der Praxis unsinnig. Jedoch wird einem dadurch nahegelegt f_I proportional zu fg zu wählen. Ist zum Beispiel g die Indikatorfunktion also

$g(x) = \mathbf{1}_{\{x \in A\}}$ wobei $A \subset \mathbb{R}$ gilt. Dann ist $\alpha = \mathbb{P}(X \in A)$ und die Importance Sampling Dichte, welche einen Schätzer mit Varianz gleich 0 liefert ist $\frac{g(x)f(x)}{\alpha}$, wenn $\alpha > 0$, was gerade die bedingte Dichte von X gegeben $X \in A$ ist. Wenn man Importance Sampling also auf die Berechnung von Wahrscheinlichkeiten anwenden will, sollte man eine Dichte suchen, welche die bedingte Dichte approximiert. Man sollte daher f_I so wählen, dass das Ereignis $\{X \in A\}$ öfter simuliert wird, insbesondere dann, wenn A eine Menge ist, welche unter f selten erreicht wird.

Dieses Vorgehen wird im folgenden Beispiel angewendet:

Beispiel 5.2. *Es wird die Wahrscheinlichkeit $\mathbb{P}(X > 3.5)$ gesucht, wobei $X \sim N(0, 1)$ ist. Da dieses Ereignis ein seltenes Ereignis ist, werden sehr viele Stichproben benötigt um einen guten Schätzer für diese Wahrscheinlichkeit zu erhalten. Hier kann die Methode Importance Sampling aushelfen. Wählt man als Importance Sampling Dichte f_I die Dichte der Normalverteilung mit Erwartungswert $\mu = 4$ und Varianz $\sigma = 1$, dann ist $f_I(x) > 0$ für $f(x) > 0$, wobei f die Dichte der Standardnormalverteilung ist. Unter der Normalverteilung mit Dichtefunktion f_I ist das Ereignis kein seltenes. Der Likelihood Ratio von f und f_I ist*

$$\frac{f(x)}{f_I(x)} = \exp\left(\frac{-2x\mu + \mu^2}{2}\right).$$

Der exakte Wert dieser Wahrscheinlichkeit ist $\mathbb{P}(X > 3.5) = 0.00023$. Es werden 5000 Stichproben zu je 100 Pseudozufallszahlen berechnet, die Tabelle 5.2 enthält die Werte mit Importance Sampling und klassische Monte Carlo Schätzung. Man kann sehen, dass die Varianz und die Genauigkeit der Schätzer mit Importance Sampling besser sind als der klassische Schätzer.

	Schätzung	Varianz
klassische Monte Carlo Schätzung	0,000236	$2,345 \cdot 10^{-6}$
Schätzung mit Importance Sampling	0,000232	$2,202 \cdot 10^{-9}$

Tabelle 2: Wahrscheinlichkeitsschätzung mittels Importance Sampling

5.4.4 Stratified Sampling

Bei dieser Methode teilt man die Grundmenge aus der die Stichprobe gezogen wird in disjunkte Teilmengen ein, und zieht aus jeder Teilmenge eine bestimmte Anzahl an Stichprobenpunkte. Damit will man Stichproben, welche sich nur auf eine Teilmenge der Grundmenge konzentrieren verhindern. Ziel ist es wiederum mit dieser Methode die Varianz zu verringern.

Es wird wieder $Y = g(Z)$ gesetzt, des weiteren sei X eine Zufallsvariable, welche zur selben Zeit wie Y generiert wird. Dann gilt

$$E[Y] = E[E[Y|X]]$$

Teilt man \mathbb{R} nun in N disjunkte Teilmengen A_1, \dots, A_N ein, so dass $\mathbb{P}\left[X \in \bigcup_{i=1}^N A_i\right] = 1$ gilt. Dann kann man den zu schätzenden Erwartungswert als

$$E[Y] = \sum_{i=1}^N \mathbb{P}[X \in A_i] E[Y|X \in A_i] = \sum_{i=1}^N p_i E[Y|X \in A_i]$$

schreiben, wobei $p_i = \mathbb{P}[X \in A_i]$ ist.

Wenn p_i bekannt oder leicht zu berechnen ist, kann der Erwartungswert von Y geschätzt werden, indem der bedingte Erwartungswert durch Monte Carlo Simulation approximiert wird. Dadurch erhält man den erwartungstreuen Schätzer

$$\hat{\alpha}_Y^S = \sum_{i=1}^N p_i \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}, \quad (35)$$

wobei Y_{ij} für $j = 1, \dots, n_i$, $i = 1, \dots, N$ unabhängige Realisationen der bedingten Verteilung von Y gegeben $X \in A_i$.

Es wird nun der Erwartungswert und die Varianz dieses Schätzers untersucht, dazu sei die Stichprobengröße mit $n \in \mathbb{N}$ fix festgelegt. Weiter sei

- $\mu = E[Y]$
- $\mu_i = E[Y_{ij}] = E[Y|X \in A_i]$
- $\sigma = Var[Y]$
- $\sigma_i = Var[Y_{ij}] = Var[Y|X \in A_i]$

Der Schätzer ist erwartungstreu, denn es gilt

$$E[\hat{\alpha}_Y^S] = \sum_{i=1}^N p_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} E[Y_{ij}] \right) = \sum_{i=1}^N p_i \mu_i = \mu.$$

Für die Berechnung der Varianz werden zwei verschiedene Fälle betrachtet. Der erste Fall betrachtet eine proportionale Stichprobenanzahl aus den verschiedenen Teilmengen, und der zweite behandelt den Fall, dass die Stichprobenanzahl aus den disjunkten Teilmengen beliebig gewählt wird.

1. Es gelte zurest die Stichprobengröße in den Teilmengen A_i sei $n_i = np_i$ für $i = 1, \dots, N$. Es wird angenommen, dass $np_i \in \mathbb{N}$ gilt. Dann gilt für die $Var[\hat{\alpha}_Y^S]$:

$$Var[\hat{\alpha}_Y^S] = \sum_{i=1}^N p_i^2 \left(\frac{1}{n_i^2} \sum_{j=1}^{n_i} Var[Y_{ij}] \right)$$

$$= \sum_{i=1}^N p_i^2 \left(\frac{1}{n_i^2} \sum_{j=1}^{n_i} \sigma_i^2 \right) = \sum_{i=1}^N p_i^2 n_i \frac{1}{n_i^2} \sigma_i^2 \stackrel{p_i = \frac{n_i}{n}}{=} \frac{1}{n} \sum_{i=1}^N p_i \sigma_i^2.$$

Da die Varianz des klassischen Monte Carlo Schätzer durch $Var[\hat{\alpha}_Y] = \frac{Var[Y]}{n}$ gegeben ist, folgt

$$\begin{aligned} Var[\hat{\alpha}_Y] &= \frac{1}{n} (E[Y^2] - E[Y]^2) \\ &= \frac{1}{n} \left(\sum_{i=1}^N p_i E[Y^2|X \in A_i] - \left(\sum_{i=1}^N p_i E[Y] \right)^2 \right) \\ &= \frac{1}{n} \left(\sum_{i=1}^N p_i (Var[Y|X \in A_i] + (E[Y|X \in A_i])^2) - \left(\sum_{i=1}^N p_i \mu_i \right)^2 \right) \\ &= \frac{1}{n} \left(\sum_{i=1}^N p_i \sigma_i^2 + \sum_{i=1}^N p_i \mu_i^2 - \left(\sum_{i=1}^N p_i \mu_i \right)^2 \right) = \quad (36) \\ &Var[\hat{\alpha}_Y^S] + \frac{1}{n} \left(\sum_{i=1}^N p_i \mu_i^2 - \left(\sum_{i=1}^N p_i \mu_i \right)^2 \right), \end{aligned}$$

da $f(x) = x^2$ eine konvexe Funktion ist und $\sum_{i=1}^N p_i = 1$ gilt, folgt mit der Jensenschen Ungleichung

$$\left(\sum_{i=1}^N p_i \mu_i \right)^2 \leq \sum_{i=1}^N p_i \mu_i^2,$$

was bedeutet, dass bei einer proportionalen Wahl der n_i auf alle Fälle keine Verschlechterung der Varianz herbeigeführt wird.

2. Wählt man die n_i für $i = 1, \dots, N$, sodass $\sum_{i=1}^N n_i = n$, beliebig, also nicht proportional zu den p_i , dann gilt für die Varianz von $\hat{\alpha}_Y^S$

$$Var[\hat{\alpha}_Y^S] = \sum_{i=1}^N \frac{p_i^2}{n_i} \sigma_i^2.$$

Was man unter der Nebenbedingung $\sum_{i=1}^N n_i = n$ leicht minimieren kann. Dazu verwendet man die Lagrange-Methode, wobei die Lagrange-funktion wie folgt gegeben ist:

$$L(n_1, \dots, n_N, \lambda) = \sum_{i=1}^N \frac{p_i^2}{n_i} \sigma_i^2 + \lambda \left(\sum_{i=1}^N n_i - n \right).$$

Differenziert man nach n_i für $i = 1, \dots, N$ und nach λ erhält man die Gleichungen

$$0 \stackrel{!}{=} \frac{d}{dn_i} L(n_1, \dots, n_N, \lambda) = -\frac{p_i^2}{n_i^2} \sigma_i^2 + \lambda, \text{ für } i = 1, \dots, N$$

$$0 \stackrel{!}{=} \frac{d}{d\lambda} L(n_1, \dots, n_N, \lambda) = \sum_{i=1}^N n_i - n$$

Die Lösungen dieser Gleichungen sind durch

$$n_i^* = n \frac{p_i \sigma_i}{\sum_{i=1}^N p_i \sigma_i}, \text{ für } i = 1, \dots, N,$$

gegeben. Durch einsetzen wie in [GNR12] sieht man hier leicht, dass die Varianz bei dieser Wahl der Stichprobenanzahl minimiert wird, den man erhält

$$\text{Var} [\hat{\alpha}_Y^S] = \frac{1}{n} \left(\sum_{i=1}^N p_i \sigma_i \right)^2.$$

Verglichen mit der Varianz im proportionalen Fall, sieht man, dass die zusätzliche Reduktion der Varianz am größten ist, wenn die Standardabweichungen der Schichten sich stark unterscheiden.

Da in der Praxis die bedingten Varianzen jedoch im Allgemeinen nicht bekannt sind, können die optimalen n_i^* nicht direkt berechnet werden. Daher werden meist aus einem kleineren Testlauf die Varianzen geschätzt, mit welchen die n_i^* dann festgelegt werden, welche dann in der Hauptsimulation angewendet werden. Für den Testlauf wird eine kleinere Stichprobenanzahl simuliert als dann in der richtigen Berechnung.

Die einfachste Anwendung dieser Methode ist die Anwendung auf gleichverteilte Zufallsvariablen, welche eine Simulation lenken. Die Vorgehensweise für die Standardgleichverteilung wird im folgenden Beispiel erläutert:

Beispiel 5.3. *Hat man U_1, \dots, U_n gleichverteilte Zufallsvariablen auf dem Einheitsintervall $(0, 1)$, dann kann man dieses Intervall in n Teilintervalle gleicher Länge aufteilen:*

$$A_1 = \left(0, \frac{1}{n}\right], A_2 = \left(\frac{1}{n}, \frac{2}{n}\right], \dots, A_n = \left(\frac{n-1}{n}, 1\right).$$

Jedes dieser Intervalle hat unter der Gleichverteilung Wahrscheinlichkeit $\frac{1}{n}$. Die Stichprobengröße und die Anzahl der Teilmengen sind hier gleich groß. Setzt man nun

$$V_i = \frac{i-1}{n} + \frac{U_i}{n} \tag{37}$$

für $i = 1, \dots, n$, dann ist V_i gleichverteilt auf dem Intervall $\frac{i-1}{n}$ und $\frac{i}{n}$. V_1, \dots, V_n ergeben somit eine stratified Stichprobe der Gleichverteilung.

Durch dieses Beispiel kann man wegen Satz 2.3, die Methode der Stratified Samples auf beliebige Verteilungen anwenden. Hier wird die Anwendung auf die Standardnormalverteilung betrachtet. Im dem Unterkapitel 5.4.4 wird diese Methode des Stratified Samples auf den Endwert einer Brownsche Bewegung erweitert.

Beispiel 5.4. *Es sei Φ die Verteilungsfunktion der Standardnormalverteilung und Φ^{-1} ihre Inverse. Es seien des weiteren p_1, \dots, p_K Wahrscheinlichkeiten, welche aufsummiert 1 ergeben. Dann definiert man $a_0 = -\infty$,*

$$a_1 = \Phi^{-1}(p_1), a_2 = \Phi^{-1}(p_1 + p_2), \dots, a_K = \Phi^{-1}(p_1 + \dots + p_K) = \Phi^{-1}(1),$$

damit kann man die reellen Zahlen in K disjunkte Teilmengen

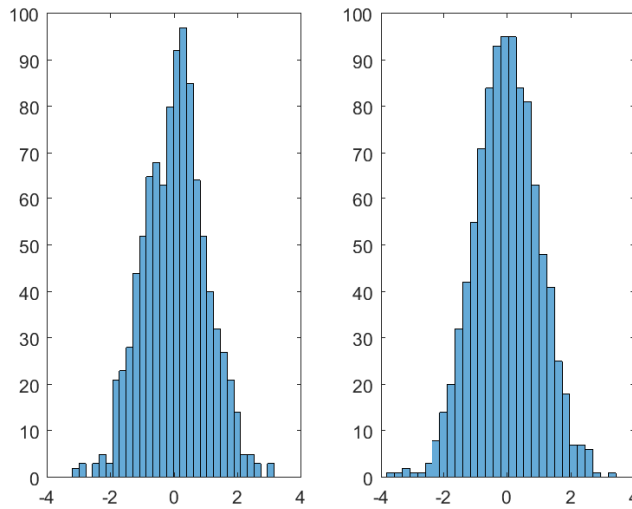
$$A_1 = (a_0, a_1], \dots, A_K = (a_{K-1}, a_K]$$

bzw. $A_K = (a_{K-1}, a_K)$, wenn $a_K = \infty$ was bei der Normalverteilung der Fall ist, einteilen. Jede Teilmenge A_i hat unter Φ Wahrscheinlichkeit p_i . Um die Teilmengen A_1, \dots, A_K für Stratified Samples zu benutzen, muss es möglich sein Y bedingt auf $Y \in A_i$ zu generieren. Dazu kann man die Inversionsmethode verwenden. Sei U eine standardgleichverteilte Zufallsvariable, dann setze

$$V = \Phi(a_{i-1}) + U(\Phi(a_i) - \Phi(a_{i-1})).$$

Es gilt, dass V gleichverteilt ist zwischen $\sum_{j=1}^{i-1} p_j$ und $\sum_{j=1}^i p_j$. Man erhält nun mit $\phi^{-1}(V)$ eine Zufallsvariable, welche die Verteilung von Y bedingt auf $Y \in A_i$ hat.

In der Graphik 1 sieht man den Unterschied zwischen 1000 standardnormalverteilten Pseudo-Zufallszahlen und durch Stratified Sampling erzeugte standardnormalverteilte Zufallszahlen mit $K = 100$ Teilmengen, wobei aus jeder Teilmenge 10 Stichproben gesammelt wurden.



Simulation von standardnormalverteilten Zufallszahlen, auf der rechten wurde Stratified Sample verwendet mit 100 Teilmengen, welche jeweils die gleiche Wahrscheinlichkeit haben. Anzahl der Stichproben ist bei beiden Histogrammen 1000, also auf der rechten Seite 10 pro Teilmenge, und bei beiden Histogrammen wurden die Stichproben auf 30 bins aufgeteilt.

Abbildung 1: Standardnormalverteilte Stichproben

Terminal Stratification Es wird nun wie vorher schon erwähnt die Methode des Stratified Samples für die Normalverteilung auf den Endwert einer Brownschen Bewegung erweitert. Die Überlegung dahinter ist, da manche Optionen stark vom Wert am Ende ihrer Laufzeit abhängen, wie es zum Beispiel bei Europäischen Optionen der Fall ist, kann man durch die Methode des Stratified Samples einen Teil der Varianz eliminieren, indem man sie auf den Endwert der Option anwendet. Bei Optionen deren Preise von einer geometrischen Brownschen Bewegung mit konstanter Volatilität abhängen, ist das gleichbedeutend damit die Methode auf den Endwert der Brownschen Bewegung, welche die geometrische lenkt, anzuwenden, da die geometrische Brownsche Bewegung eine monotone Transformation der Brownschen Bewegung ist.

Man will Pfade einer diskreten Brownschen Bewegung $W(t_1), \dots, W(t_m)$ simulieren, wobei der Endwert in eine bestimmte Teilmenge der reellen Zahlen fällt. Dazu kann man die Methode der Brownschen Brücke verwenden. Da der Endwert der Brownschen Bewegung normalverteilt ist mit Erwartungswert 0 und Varianz t_m , kann dieser zuerst wie im Beispiel 5.4 simuliert werden, dann kann der restliche Pfad durch eine Brownsche Brücke bedingt auf den Endwert erzeugt werden.

Betrachte man nun K Teilmengen A_1, \dots, A_K , welche alle die gleiche Wahrscheinlichkeit $\frac{1}{K}$ haben. Die Anzahl der Stichproben ist in allen Teilmengen gleich, hier wird nur eine Pseudo-Zufallszahl aus jeder Teilmenge genommen. Seien nun U_1, \dots, U_K unabhängige standardgleichverteilte Zufallsvariablen und man setzt

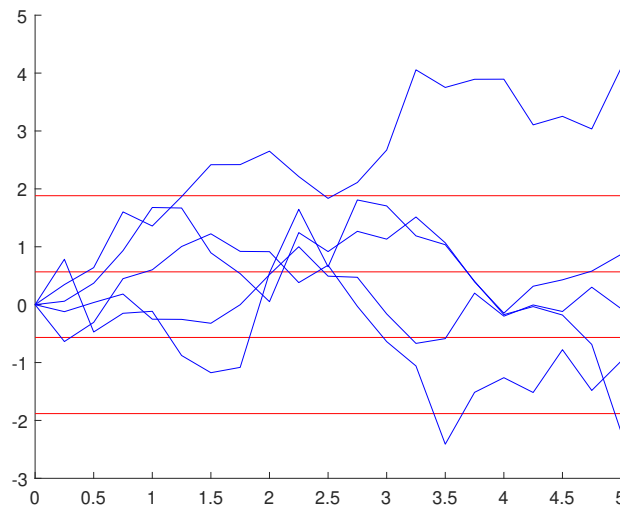
$$V_i = \frac{i-1}{K} + \frac{U_i}{K}.$$

Dann bilden $\Phi^{-1}(V_1), \dots, \Phi^{-1}(V_K)$ eine standardnormalverteilte Stichprobe. Die Endwerte der K Pfade der Brownschen Bewegung erhält man, wenn man $W_1(t_m) = \sqrt{t_m}\Phi^{-1}(V_1), \dots, W_K = \sqrt{t_m}\Phi^{-1}(V_K)$ setzt. Um die Pfade aufzufüllen, kann man nun die bedingte Verteilung verwenden, über welche man auch eine Brownsche Brücke konstruiert. Für $i = 1, \dots, m-1$ kann man rekursiv $W(t_i)$ durch

$$W(t_i) = \frac{t_m - t_i}{t_m - t_{i-1}} W(t_{i-1}) + \frac{t_i - t_{i-1}}{t_m - t_{i-1}} W(t_m) + \sqrt{\frac{(t_m - t_i)(t_i - t_{i-1})}{t_m - t_{i-1}}} Z_i$$

bestimmen, wobei $W(0) = 0$ und die Z_i für $i = 1, \dots, m-1$ sind unabhängige standardnormalverteilte Zufallsvariablen.

In Abbildung 2 sieht man 5 Pfade einer Brownschen Bewegung, welche jeweils in einer Teilmenge A_i für $i = 1, \dots, 5$ enden.



Simulation von 5 Pfaden einer Brownschen Bewegung mit Terminal Stratification. Die Teilmengen haben alle die gleiche Wahrscheinlichkeit $p_i = 0.2$ für $i = 1, \dots, 5$. Die roten Linien zeigen die Grenzen der Teilmengen an.

Abbildung 2: Terminal Stratification Brownsche Bewegung

Zerlegung der Varianz Es wird untersucht, welcher Teil der Varianz durch die Stratified Sample Methode eliminiert wird. Es seien A_1, \dots, A_N die festgelegten Teilmengen. Es sei $\eta := \eta(X) \in \{1, \dots, N\}$ der Index der Teilmenge, in welcher X liegt ($X \in A_\eta$). Es gilt

$$Y = E[Y | \eta] + \epsilon. \quad (38)$$

Da $E[\epsilon | \eta] = E[Y - E[Y | \eta]] = 0$ und

$$E[\epsilon(E[Y | \eta] - E[Y])] = E[E[\epsilon(E[Y | \eta] - E[Y]) | \eta]] = 0,$$

folgt, dass ϵ und $E[Y | \eta]$ unkorreliert sind. Daher gilt

$$\text{Var}[Y] = \text{Var}[E[Y | \eta]] + \text{Var}[\epsilon]. \quad (39)$$

Stratified Sampling mit proportionaler Wahl der n_i eliminiert den ersten Term auf der rechten Seite, was eine Varianzreduktion garantiert. Was daraus folgt, dass die Varianz des Residuums gleich $E[\epsilon^2]$ ist, da $E[\epsilon] = E[Y - E[Y | \eta]] = 0$ und mit

$$E[\epsilon^2 | \eta] = E[(Y - E[Y | \eta])^2 | \eta]$$

$$E[Y^2 | \eta] - (E[Y | \eta])^2 = \text{Var}[Y | \eta]$$

bekommt man

$$\text{Var}[Y] = \text{Var}[E[Y | \eta]] + E[\text{Var}[Y | \eta]]. \quad (40)$$

$E[Y | \eta] = \mu_i$ für $\eta = i$, und die Wahrscheinlichkeit $\eta = i$ ist durch p_i geben. Damit erhält man für den ersten Teil

$$\text{Var}[E[Y | \eta]] = E[E[Y | \eta]^2] - (E[E[Y | \eta]])^2 = \sum_{i=1}^N p_i \mu_i^2 - \left(\sum_{i=1}^N p_i \mu_i \right)^2,$$

womit unter Verwendung von (39), (40) und (36) für die Varianz von ϵ

$$\text{Var}[\epsilon] = E[\text{Var}[Y | \eta]] = \sum_{i=1}^N \sigma_i^2 p_i$$

folgt. Das ist $n \text{Var}[\hat{\alpha}_Y^S]$, wenn die $n_i = np_i$ proportional gewählt werden.

Nehme man nun an, dass A_1, \dots, A_N frei wählbar ist. Die Varianz in (39) ist konstant, also gilt, dass die Residuen Varianz klein zu machen äquivalent dazu ist $\text{Var}[E[Y | \eta]]$ zu vergrößern. Man sollte also versuchen die Teilmengen A_1, \dots, A_N so zu wählen, dass die Varianz der Erwartungswerte zwischen den Teilmengen groß und die Varianz in jeder Teilmenge klein ist. Stratified Sampling eliminiert die Varianz zwischen den Schichten und die Varianz in den einzelnen Schichten bleibt. Es gilt auch, dass bei mehrmaligen anwenden dieser Methode die Varianz immer weiter reduziert werden kann.

Stratified Sampling vs. kontrollierende Zufallsvariable Um die Varianzreduktion von Stratified Sampling mit der Methode der kontrollierenden Zufallsvariable zu vergleichen, lässt man man die Anzahl der Teilmengen, welche für das Stratified Sampling verwendet werden beliebig groß werden, dann nähert sich $E[Y|\eta]$ dem $E[Y|X]$ an. Dann kann man (38)

$$Y = g(X) + \epsilon$$

schreiben. Ist die Funktion g linear, dann verkleinert die Anwendung von Stratified Sampling mit einer sehr großer Anzahl an Teilmengen die Varianz genau gleich wie wenn X als Kontrollvariable verwendet wird. Ist jedoch g nicht linear, reduziert X als Kontrollvariable nur den Teil der Varianz, welcher linear zu X ist. Bei Stratified Sampling reduziert sich die Varianz auf die Varianz des Residuums, also wird die Varianz von g gleich 0. Daher ist es effektiver die Methode des Stratified Samples anzuwenden, jedoch muss hier Vorausgesetzt werden, dass die Verteilung von X bekannt ist, und die Simulation von bedingten Stichproben sollte möglich sein. Verwendet man X als Kontrolle muss man nur den Erwartungswert von X kennen. Ist die Simulation von Y bedingt auf X sehr aufwendig, kann man folgenden Ansatz verwenden, der Poststratification genannt wird.

Poststratification Manchmal ist es sehr aufwendig bedingte Stichproben zu simulieren, obwohl die Teilmengen mit ihren Wahrscheinlichkeiten leicht zu berechnen sind. Poststratification kombiniert das Wissen der Wahrscheinlichkeiten der Teilmengen mit dem generieren herkömmlicher unabhängiger Stichproben.

Man generiere unabhängige Kopien $(X_1, Y_1), \dots, (X_n, Y_n)$ des Paares (X, Y) . Die Wahrscheinlichkeiten $p_i = \mathbb{P}(X \in A_i)$ sind für die Teilmengen A_1, \dots, A_K bekannt. Es gilt $p_i \geq 0$ für $i = 1, \dots, K$ und $\sum_{i=1}^K p_i = 1$. Dann definiere man durch

$$N_i = \sum_{j=1}^n \mathbf{1}_{\{X_j \in A_i\}}$$

die Anzahl der Stichproben von X , welche in der Teilmenge i liegen. Des weiteren sei

$$S_i = \sum_{j=1}^n \mathbf{1}_{\{X_j \in A_i\}} Y_j$$

Summe jener Y , für welche X in der Menge A_i liegt für $i = 1, \dots, K$. Das Stichprobenmittel kann dann geschrieben werden als

$$\hat{\alpha}_Y = \frac{S_1 + \dots + S_K}{n} = \sum_{i=1}^K \frac{N_i}{n} \frac{S_i}{N_i},$$

wenn alle $N_i \neq 0$. Nach dem starken Gesetz der großen Zahlen konvergiert $\frac{N_i}{n} \rightarrow p_i$ und $\frac{S_i}{N_i} \rightarrow \mu_i$ mit Wahrscheinlichkeit 1, wobei μ_i für $i = 1, \dots, K$ der

Erwartungswert von Y für $X \in A_i$ ist. Bei der Poststratification ersetzt man den Bruch $\frac{N_i}{n}$ der von den Zufallsvariablen abhängt, durch seine Erwartung p_i und man erhält den Schätzer

$$\hat{\alpha}_Y^{PS} = \sum_{i=1}^K p_i \frac{S_i}{N_i}.$$

Um die Möglichkeit abzudecken, dass keine der Beobachtungen in die Teilmenge A_i fällt, wird für diese Situation $\frac{S_i}{N_i} = 0$ gesetzt.

Aus der Konvergenz von $\frac{S_i}{N_i} \rightarrow \mu_i$ folgt, dass der Poststratification Schätzer ein Konsistenter Schätzer für den Erwartungswert ist. Diese Methode reduziert die Varianz bei wachsender Stichprobe genauso effektiv wie Stratified Sampling, jedoch kann man im Vorhinein nicht exakt sagen, ab was für einer Größer der Stichprobe die Reduktion der Varianz genauso groß ist wie beim Stratified Sampling, daher ist Stratified Sampling immer zu bevorzugen und Poststratification nur anzuwenden, wenn das Erzeugen der bedingten Stichprobe sehr aufwendig ist.

5.4.5 Latin Hypercube Sampling

Die Methode des Latin Hypercube Samplings ist eine Erweiterung des Stratified Samples für höhere Dimensionen. Hat man ein Problem höherer Dimension zu lösen, kann die Anzahl der Stichproben bei Stratified Sampling sehr groß werden, wenn man aus jeder Teilmenge zumindest eine Stichprobe sammeln will. Da man bei einem d -dimensionalen Problem, bei dem jede Dimension in K Teilmengen eingeteilt wird, zumindest K^d Stichproben sammeln muss um aus jeder Teilmenge eine Stichprobe zu erhalten.

Latin Hypercube Sampling ist eine Erweiterung von Stratified Sampling, welches das exponentielle Anwachsen des Stichprobenumfangs in Bezug auf die Dimension dadurch verhindert, dass alle Dimensionen als gleich angesehen werden. Die Stichproben werden so gesammelt, dass in den unterschiedlichen Dimensionen je ein Stratified Sample generiert wird. Also wird aus jeder Dimension unabhängig von den anderen simuliert, damit hat man für K Teilmengen eines d -dimensionalen Problems dK Stichproben zu generieren.

Am Besten wird diese Methode anhand des d -dimensionalen Einheitswürfels erläutert. Es sei K die fixe Stichprobenanzahl. Es wird für jede Dimension unabhängig ein Stratified Sample $V_i^{(1)}, \dots, V_i^{(K)}$ für $i = 1, \dots, d$ auf dem Einheitsintervall gesammelt, wobei die Wahrscheinlichkeit, dass die Stichprobe in die Teilmenge A_j fällt für alle $j = 1, \dots, K$ gleich ist. Dann ist jedes $V_i^{(j)}$ gleichverteilt über dem Intervall $\left[\frac{j-1}{K}, \frac{j}{K}\right)$. Schreibt man die Stratified

Samples in Spalten

$$\begin{array}{cccc} V_1^{(1)} & V_2^{(1)} & \cdots & V_d^{(1)} \\ V_1^{(2)} & V_2^{(2)} & \cdots & V_d^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ V_1^{(K)} & V_2^{(K)} & \cdots & V_d^{(K)} \end{array}$$

dann gibt jede Zeile die Koordinate eines Punktes in $[0, 1)^d$ an. Die erste Reihe ist ein Punkt im Teilwürfel $[0, \frac{1}{K})^d$, die zweite Reihe identifiziert einen Punkt in der Teilmenge $[\frac{1}{K}, \frac{2}{K})^d$ u.s.w.. Jeder dieser Punkte fällt in einen der K^d Teilwürfel, welcher entlang der Diagonale des Einheitswürfels liegt. Um daraus eine gleichverteilte Stichprobe auf dem Einheitswürfel zu erhalten, seien π_1, \dots, π_d Zufallspermutationen von $\{1, \dots, K\}$, wobei $\pi_i(j)$ gibt den Wert von j an, welcher nach der Permutation i angenommen wird. Die Reihe der Felder

$$\begin{array}{cccc} V_1^{(\pi_1(1))} & V_2^{(\pi_2(1))} & \cdots & V_d^{(\pi_d(1))} \\ V_1^{(\pi_1(2))} & V_2^{(\pi_2(2))} & \cdots & V_d^{(\pi_d(2))} \\ \vdots & \vdots & \ddots & \vdots \\ V_1^{(\pi_1(K))} & V_2^{(\pi_2(K))} & \cdots & V_d^{(\pi_d(K))} \end{array}$$

identifizieren auch weiterhin Punkte in $[0, 1)^d$, aber sie sind nicht mehr auf die Diagonale beschränkt. Es ist jetzt jeder Punkt ein Punkt, welcher auf dem d -dimensionalen Einheitswürfel gleichverteilt ist.

Um ein Latin Hypercube Sample der Größe K zu simulieren, seien $U_i^{(j)}$ unabhängige standardgleichverteilte Zufallsvariablen für $i = 1, \dots, d$ und $j = 1, \dots, K$. Seien des weiteren π_1, \dots, π_d unabhängige Zufallspermutationen von $\{1, \dots, K\}$ und setze

$$V_i^{(j)} = \frac{\pi_i(j) - 1 + U_i^{(j)}}{K} \quad (41)$$

für $i = 1, \dots, d$ und $j = 1, \dots, K$. Die Stichprobe des Latin Hypercube Samples besteht nun aus den K Punkten $(V_1^{(j)}, \dots, V_d^{(j)})$ für $j = 1, \dots, K$. Durch Anwendung der Inversionsmethode lässt sich das Latin Hypercube Sampling auf andere Verteilungen übertragen.

Beispiel 5.5. Will man eine normalverteilte Stichprobe der Größe K aus dem \mathbb{R}^d mit dem Latin Hypercube generieren, setzt man einfach für $i = 1, \dots, d$ und $j = 1, \dots, K$

$$Z_i^{(j)} = \Phi^{-1}(V_i^{(j)}) \quad (42)$$

wobei Φ^{-1} die Inverse der Verteilungsfunktion der Standardnormalverteilung ist, und die $V_i^{(j)}$ das Latin Hypercube Sample (41) ist. Die Stichprobe besteht nun aus den Vektoren $Z^{(j)} = (Z_1^{(j)}, \dots, Z_d^{(j)}) \in \mathbb{R}^d$ für $j = 1, \dots, K$.

Über die normalverteilte Stichprobe kann man auch K Pfade einer Brownschen Bewegung generieren, welche verwendet werden können um einen Prozess zu erzeugen, der von einer einzelnen Brownschen Bewegung gelenkt wird.

Beispiel 5.6. Für die Simulation von K Pfaden einer diskreten Brownschen Bewegung zu den Zeitpunkten $0 = t_0 < t_1 < \dots < t_d$, generiere man $Z^{(1)}, \dots, Z^{(K)}$ wie zuvor die Stichprobe (42). Aus diesen K Punkten aus dem \mathbb{R}^d können K diskrete Brownsche Pfade $W^{(1)}, \dots, W^{(K)}$ erzeugt werden, dazu setzt

$$W_{t_n}^{(j)} = \sum_{i=1}^n \sqrt{t_i - t_{i-1}} Z_i^{(j)} \quad (43)$$

für $n = 1, \dots, d$ und $j = 1, \dots, K$.

Dieses Latin Hypercube Sample kann man bei der Simulation von pfadabhängigen Optionen wie zum Beispiel Asiatische Optionen einsetzen.

Varianzreduktion Um die Untersuchung der Varianz zu vereinfachen, nehme man an, dass das Ziel die Schätzung von

$$\alpha_f = \int_{[0,1]^d} f(u) du$$

ist, wobei $f : [0, 1]^d \rightarrow \mathbb{R}$ eine quadratisch integrierbare Funktion ist. Der klassische Monte Carlo Schätzer dieses Integrals kann als

$$\bar{\alpha}_f = \frac{1}{K} \sum_{j=1}^{K-1} f(U_{jd+1}, \dots, U_{jd+d})$$

geschrieben werden, wobei U_1, U_2, \dots unabhängige standardgleichverteilte Zufallsvariablen sind. Die Varianz dieses Schätzers ist $\frac{\sigma^2}{K}$ mit $\sigma^2 = \text{Var}[f(U_1, \dots, U_d)]$. für das Latin Hypercube Sample $V^{(1)}, \dots, V^{(K)}$, welches wie in (41) berechnet wird, ist der Schätzer definiert durch

$$\hat{\alpha}_f = \frac{1}{K} \sum_{j=1}^K f(V^{(j)}).$$

Für die Varianz dieses Schätzers gilt

$$\text{Var}[\hat{\alpha}_f] \leq \frac{\sigma^2}{K-1}$$

für $K \geq 2$, was genauer in [Gla04] erläutert wird. Damit erhält man bei der Verwendung eines Latin Hypercube Samples, dass die Varianz nicht größer ist, als jene einer unabhängigen Stichprobe der Größe $K-1$.

6 Monte Carlo Methoden in der Versicherungsmathematik

Ziel ist es eine Methode zu finden um die Eigenmittelkapitalanforderung für ein Versicherungsunternehmen bezüglich dem Value at Risk und dem Expected Shortfall schätzen zu können. Um die Risikomaße zu berechnen, benötigt man die Verteilung des zukünftigen Wertes des Portfolios. Da es für viele Positionen in einem Portfolio keine analytische Bepreisungsformel gibt, ist es für ein gutes Modell, welches für Berechnung der Eigenmittelanforderungen eines Versicherungsunternehmens benutzt wird, nötig einen auf Simulationen basierenden Algorithmus zu verwenden. Im vorigen Kapitel wurde so eine Methode, nämlich die Monte Carlo Simulation vorgestellt. In diesem Kapitel werden verschiedene Ansätze der Monte Carlo Methode, wie sie in der Finanz- und Versicherungsmathematik zum Einsatz kommt, betrachtet und miteinander verglichen.

Eine naive Herangehensweise legt die Nested Monte Carlo Simulation nahe, welche zuerst kurz beschrieben wird. Diese Methode ist aber nicht sehr effizient, daher werden dann im weiteren der Least Square Monte Carlo Ansatz und die Methode der replizierenden Portfolios genauer betrachtet.

6.1 Nested Monte Carlo Simulationen

Der Inhalt dieses Kapitels orientiert sich an dem Paper [BBR10] und an dem Artikel [GJ10].

Die Nested Simulationen sind geschachtelte Monte Carlo Simulationen. In diesem Fall wird im äußeren Schritt jeweils eine Stichprobe der Risikofaktoren durch eine Monte Carlo Simulation über den Risikohorizont $[0, t]$ generiert, und im inneren Schritt wird jede Position des Portfolios in den Realisationen der Risikofaktoren über den Projektionshorizont neu bepreist. Aus dem daraus resultierenden diskontierten Verlusten zur Zeit t kann die Eigenmittelkapitalanforderung durch Anwenden eines Risikomaßes berechnet werden. Der Rechenaufwand vor allem in der Neubepreisung ist enorm.

Die diskontierte Auszahlung bzw. die Verluste zum Zeitpunkt t wird als $L_t = E_{\mathbb{Q}}[X|\mathcal{F}_t]$ dargestellt. In dieser Darstellung kann die Filtration durch $(G_t, A_t(G))$ ersetzt werden, da $A_t(G)$ alle für die Auszahlungsfunktion relevanten Informationen beinhaltet.

Für den äußeren Schritt der Nested Simulationen werden $N \in \mathbb{N}$ Pfade $(G_s^i)_{s \in [0, t]}$ mit dazugehörigen Zustandsvariablen $(G_t^i, A_t^i(G))$, $i \in \{1, \dots, N\}$, unter dem realen Maß \mathbb{P} generiert. Der zukünftige Zahlungsstrom bedingt auf den Zustand des Finanzmarktes im Szenario i kann dann durch

$$L_t^i = E_{\mathbb{Q}}[X|(G_t^i, A_t^i(G))] \quad (44)$$

beschrieben werden.

Im Allgemeinen kann (44) jedoch nicht analytisch gelöst werden. Die Berechnung erfolgt mit einer weiteren Monte Carlo Simulation im sogenannten inneren Schritt. Basierend auf den Zustand $(G_t^i, A_t^i(G))$ in den „real-world“ Szenario $i \in \{1, \dots, N\}$, werden $K^i \in \mathbb{N}$ risikoneutrale Szenarien simuliert und mit $(G_s^{(i,k)})_{s \in [t, T]}$ bezeichnet. Damit kann man die Auszahlung $X^{(i,k)}$ für $k = 1, \dots, K^i$ und $i = 1, \dots, N$ zum Zeitpunkt T berechnen. Einen MC-Schätzer für L_t^i bekommt man nun durch

$$L_t^i = \frac{1}{K^i} \sum_{k=1}^{K^i} X^{(i,k)}, \quad i \in \{1, \dots, N\}.$$

Wobei man nun daraus leicht eine Schätzung für die Risikomaße Expected Shortfall oder Value at Risk berechnen kann. Diese Methode wird in der Abbildung 3 dargestellt.

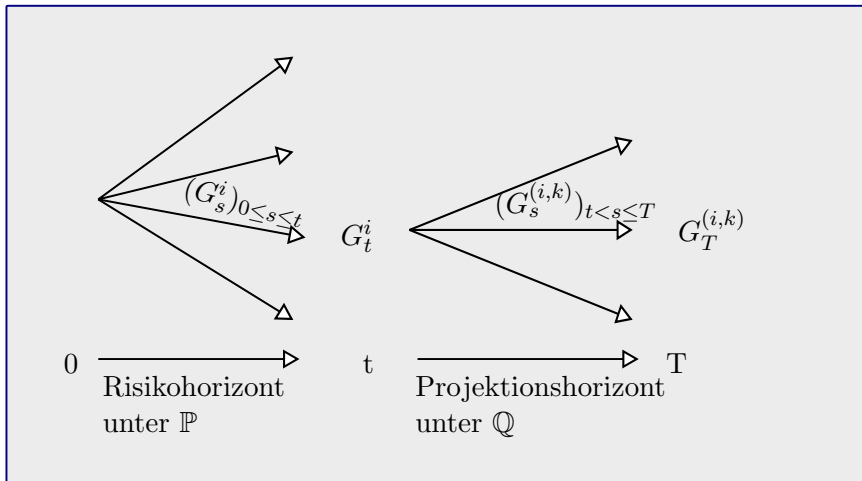


Abbildung 3: Nested Simulation

Während es im äußeren Schritt eine Abhängigkeit über die Positionen gibt, können die Simulationen des inneren Schritts unabhängig von einander durchgeführt werden, da der Wert der Position i zur Zeit t ein einfacher bedingter Erwartungswert der eigenen nachfolgenden Zahlungsströme ist und nicht von den zukünftigen Zahlungsströmen der anderen Positionen abhängt.

Es wäre daher effektiver die Simulation der inneren Schritte gleichzeitig über alle Positionen laufen zu lassen um in Folge die Anzahl der generierten Pfade von G_s auf $(t, T]$ zu reduzieren. Jedoch geht die Unabhängigkeit der Bepreisungsfehler, welche deterministische Funktionen dieser simulierten Pfade sind, verloren, wenn man die Stichproben von G_s nicht separat für jede Position generiert. Was einen Nachteil bei der Neupebreisung der

einzelnen Positionen nach sich zieht. Dieses Vorgehen ist daher nicht empfehlenswert.

6.2 Least Square Monte Carlo und replizierendes Portfolio

Da der rechnerische und zeitliche Aufwand der Nested Monte Carlo Simulationen für die Lebensversicherung zu groß ist, hat man Methoden zur Berechnung der Eigenmittelanforderung entwickelt, welche Approximationsmethoden mit Monte Carlo Techniken kombinieren. Es werden die Ansätze Least Square Monte Carlo (LSMC) und replizierendes Portfolio (RP) genauer betrachtet. Diese Methoden zielen darauf ab die Anzahl der inneren Schritte zu minimieren. Verwendet man eine dieser beiden Methoden verringert dies den Aufwand für die Risikokalkulation enorm. Beim Vergleich dieser zwei Methoden konzentriert man sich hier auf den Regressionsfehler, der den Unterschied der beiden Methoden ausmacht. Bei diesen Betrachtungen wird sich herausstellen, dass das replizierende Portfolio zu bevorzugen ist, auch wenn die Berechnung etwas aufwändiger sein kann. Die Arbeit [PS15] wird als Grundlage für dieses Kapitel verwendet.

Die LSMC Methode ergibt eine Approximationsfunktion für die bedingte Erwartungsfunktion zur Zeit t . Daraus kann man die empirische Verteilung zu dieser Zeit herleiten, aus der das Risikokapital berechnet werden kann. Dieser Ansatz wird in 4 illustriert.

Das Prinzip der Replikation ist es ein Portfolio von Finanzinstrumenten zu konstruieren, welches die Auszahlungsfunktion der Zielzufallsvariable zur Zeit T widerspiegelt. Hat man ein RP bestehend aus Instrumenten, für welche man die Werte kennt, kann daraus der t -Wert und daher auch die empirische Verteilung schnell berechnet werden, wodurch wieder die Berechnung des Risikokapitals ermöglicht wird, was in 5 illustriert wird.

Sowohl in LSMC als auch in RP wird die Approximation durch eine Least Square Methode gewonnen, jedoch unterscheiden sich die Datenpunkte. Es werden nun die Least Square Ansätze zur Konstruktion der LSMC und RP Schätzer erläutert.

6.2.1 Regressionsmodell für LSMC

In der LSMC Methode wird die Schätzung in einem Schritt erhalten, indem man für eine Stichprobe der Nested Monte Carlo Simulation mit nur einem inneren Schritt geeignete Basisfunktionen durch die Least Squares Methode findet, um den Wert der bedingte Erwartungsfunktion zur Zeit t möglichst genau zu approximieren.

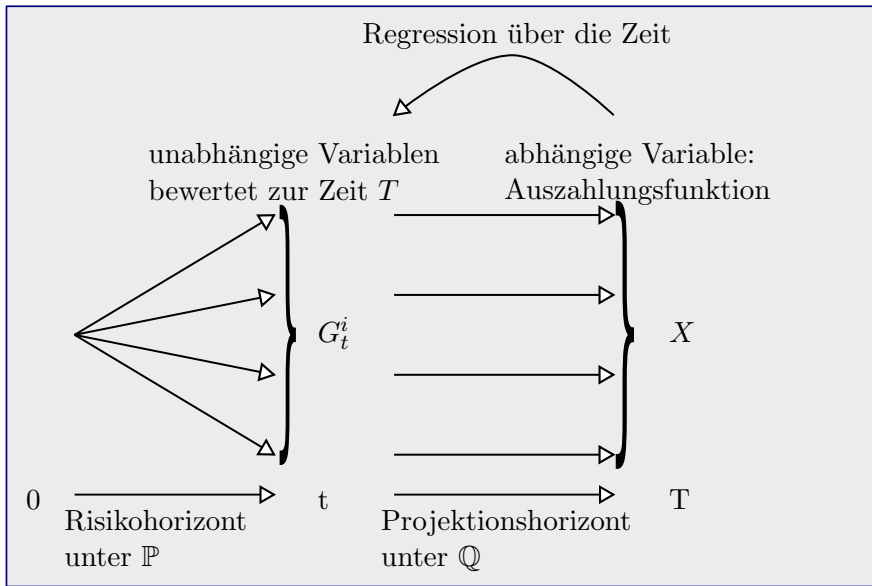


Abbildung 4: Least Square Monte Carlo

Um den LSMC Ansatz zu erläutern wird angenommen, dass sich der bedingte Erwartungswert schreiben lässt als

$$g_{0,t}(A_t(G)) = L_t = E_{\mathbb{Q}}[X|\mathcal{F}_t], \quad (45)$$

wobei $A_t : \mathbb{D}_d[0, t] \rightarrow \mathbb{R}^{l_t}$ eine bekannte messbare Funktion ist und $g_{0,t} : \mathbb{R}^{l_t} \rightarrow \mathbb{R}$ eine unbekannte Borel-messbar Funktion. Die tiefgestellte 0 wird verwendet um die bedingte Erwartungsfunktion von der Auszahlungsfunktion $g_T(A_T(G))$, die ja bekannt ist, zu unterscheiden.

Da X quadratisch integrierbar ist, ist auch $E_{\mathbb{Q}}[X|\mathcal{F}_t]$ quadratisch integrierbar, woraus folgt, dass $g_{0,t} \in L^2(\mathbb{Q})$. Dieser Raum ist ein separabler Hilbertraum, das heißt er besitzt eine abzählbare orthonormale Basis. Daher kann man

$$g_{0,t} = \sum_{k=1}^{\infty} \beta_k v_k$$

schreiben, wobei $\{v_k\}_{k=1}^{\infty}$ eine abzählbare orthonormale Basis des $L^2(\mathbb{Q})$ ist. Wie in Korollar 2.2 gezeigt wurde, kann der bedingte Erwartungswert auch als orthogonale Projektion gesehen werden, daher sind die Koeffizienten für die Berechnung von $g_{0,t}$ durch

$$\beta_k = E_{\mathbb{Q}}[X v_k(A_t(G))]. \quad (46)$$

gegeben. Somit gilt nun speziell für (45)

$$g_{0,t}(A_t(G)) = \sum_{k=1}^{\infty} \beta_k v_k(A_t(G)), \quad (47)$$

wodurch ein Projektionsfehler entsteht, welcher wie folgt definiert wird:

$$p_{0,t}(A_T(G)) = X - g_{0,t}(A_t(G)). \quad (48)$$

Im LSMC Ansatz wird versucht die unbekannte Funktion $g_{0,t}$ durch die Darstellung (47) zu berechnen, wobei hier von der Simulation der Daten unter dem risikoneutralen Maß \mathbb{Q} ausgegangen wird. Die Anzahl der unbekanntenen Koeffizienten kann jedoch unendlich sein, daher ist die Gleichung nicht direkt lösbar, sondern man muss die Basisfunktionen auf eine endliche Anzahl $K < \infty$ reduzieren, dass heißt es wird $\{v_k\}_{k=1}^K$ verwendet, was zu

$$g_{0,t}^K = \sum_{k=1}^K \beta_k v_k = (\beta^K)^T v^K \quad (49)$$

für $\beta^K = (\beta_1, \dots, \beta_K)^T$ und $v^K = (v_1, \dots, v_K)^T$ Führt. Dadurch erhält man einen Approximationsfehler

$$a_{0,t}^K = g_{0,t} - g_{0,t}^K. \quad (50)$$

Man beachte, dass

$$E_{\mathbb{Q}} [g_{0,t}^K(A_t(G)) a_{0,t}^K(A_t(G))] = E_{\mathbb{Q}} \left[\sum_{k=1}^K \beta_k v_k(A_t(G)) \left(\sum_{k=1}^{\infty} \beta_k v_k(A_t(G)) - \sum_{k=1}^K \beta_k v_k(A_t(G)) \right) \right] = 0$$

da die Basis $\{v_k\}_{k=1}^{\infty}$ orthonormal ist. Für $K \rightarrow \infty$ konvergiert der Approximationsfehler gegen 0.

Nimmt man alles zusammen, kann X durch folgende Regressionsgleichung ausgedrückt werden

$$X = g_{0,t}^K(A_t(G)) + a_{0,t}^K(A_t(G)) + p_{0,t}(A_T(G)), \quad (51)$$

wobei die Summe von Approximationsfehler und Projektionsfehler den Regressionsfehler ergibt.

Da X und G nicht direkt berechenbar sind, müssen sie durch eine Monte Carlo Simulation generiert werden. Sei nun $((X_1, A_t(G_1)), \dots, (X_n, A_t(G_n)))$ eine simulierte Stichprobe der Größe n . $g_{0,t}^K$ wird dann durch das Optimierungsproblem

$$\hat{g}_{0,t}^K = \arg \min_{g \in \mathcal{H}_K} \frac{1}{n} \sum_{i=1}^n (X_i - g(A_t(G_i)))^2,$$

wobei $\mathcal{H}_K = \left\{ g : \mathbb{R}^d \rightarrow \mathbb{R} \mid g = \sum_{k=1}^K \beta_k v_k, \beta_k \in \mathbb{R} \right\}$, geschätzt. Die Lösung dieses Minimierungsproblems ist durch

$$\hat{g}_{0,t}^K = (\hat{\beta}^K)^T v^K \quad (52)$$

gegeben, wobei

$$\hat{\beta}^K = ((V_K)^T V_K)^{-1} (V_K)^T \hat{X}$$

mit $\hat{X} = (X^1, \dots, X^n)^T$ und V_K ist eine $(n \times K)$ -Matrix, deren Reihen gleich den Vektoren $v^K(A_t(G_i))$ für $i = 1, \dots, n$ sind.

Der LSMC Schätzer wird durch eine Regression der Auszahlungsfunktion X konstruiert, wobei die unabhängigen Variablen zur Zeit t bewertet werden. Dadurch erhält man natürlich auch einen Schätzer für die bedingte Erwartungsfunktion $E_{\mathbb{Q}}[X | \mathcal{F}_t]$.

Dieser Schätzer ergibt einen Fehler, welcher sich aus dem Approximations-, Projektions- und Schätzungsfehler ergibt. Der Schätzungsfehler wird auch als Monte Carlo Fehler bezeichnet. Während der Approximations- für $K \rightarrow \infty$ und der Monte Carlo Fehler für $n \rightarrow \infty$ gegen 0 strebt, kann der Projektionsfehler nur eliminiert werden, wenn man die Auszahlungsfunktion X , bewertet zur Zeit T , durch Basisfunktionen, welche zum selben Zeitpunkt bewertet werden, berechnet, was im Prinzip der Ansatz des replizierenden Portfolio ist.

6.2.2 Regressionsmodell für RP

In diesem Unterkapitel wird das Modell des replizierenden Portfolios erläutert. Im Gegensatz zur LSMC Methode, wo die bedingte Erwartungsfunktion direkt geschätzt wird, wird beim RP Ansatz zuerst eine Schätzung für die Auszahlungsfunktion X konstruiert, das heißt es wird ein replizierendes Portfolio aus einfachen Finanzinstrumenten für X erstellt. Hat man eine lineare Repräsentation durch Basisfunktionen, welche Funktionen der einfachen Finanzinstrumente sind, gefunden, dann wird der bedingte Erwartungsoperator $E_{\mathbb{Q}}[\cdot | \mathcal{F}_t]$ auf diese angewendet. Damit bekommt man den gesuchten Wert des Portfolios zum Zeitpunkt t . Dieser Ansatz, welcher zwei Schritte beinhaltet, ist nur sinnvoll, wenn für die Auszahlungsfunktion X Basisfunktionen verwendet werden, für welche die bedingte Erwartung leicht berechnet werden kann. Dies impliziert für das risikoneutrale Maß, dass bereits eine analytische Lösungsformel für den Preis jener Finanzinstrumente

verfügbar ist, welche als Basisfunktionen eingesetzt werden.

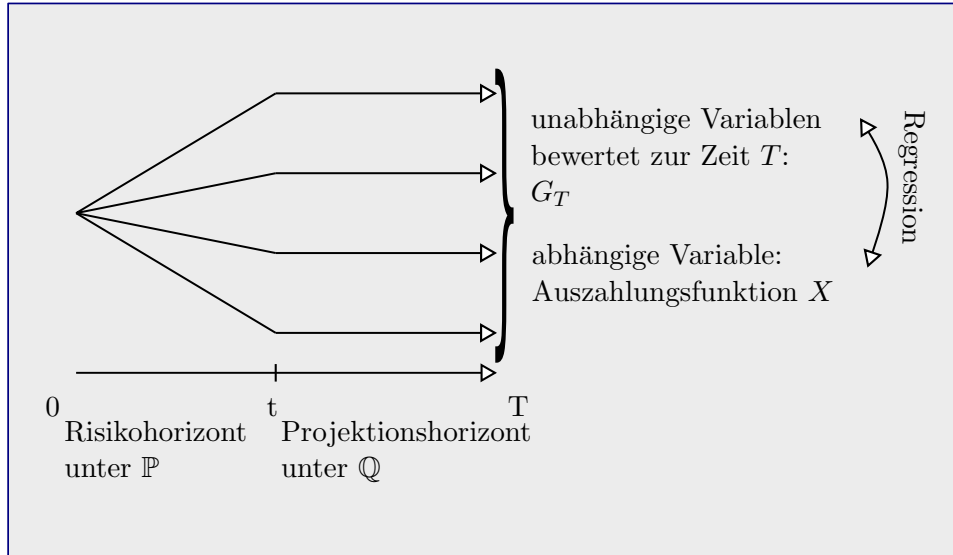


Abbildung 5: Replizierendes Portfolio

Da die Auszahlungsfunktion quadratisch integrierbar ist, kann gleich wie im Kapitel zuvor gefolgert werden, dass

$$X = g_T(A_T(G)) = \sum_{k=1}^{\infty} \alpha_k e_k(A_T(G)) \quad (53)$$

mit $\{e_k\}_{k=1}^{\infty}$ eine abzählbare orthonormale Basis des $L^2(\mathbb{Q})$ ist.

Es wird eine differenzierende Notation verwendet, da sich die Basisfunktionen für LSMC und RP unterscheiden können, wie zum Beispiel bei einer Call Option, wo die Auszahlung einen Knick hat, aber die Bedingte Erwartungsfunktion glatt ist. Daher würden sich für die Basisfunktionen in LSMC Polynome anbieten. Im Gegensatz dazu würde man sich für stückweise lineare Funktionen als Basisfunktionen für den RP-Ansatz entscheiden.

Die Koeffizienten von α sind durch

$$\alpha_k = E_{\mathbb{Q}}[X e_k(A_T(G))] \quad (54)$$

gegeben. Auch hier ist eine direkte Berechnung aufgrund der unendlichen Anzahl der Parameter nicht möglich, und die Größe der Basis muss wieder

auf eine Anzahl $K < \infty$ reduziert werden. Daher gilt

$$g_T^K = \sum_{k=1}^K \alpha_k e_k = (\alpha^K)^T e^K \quad (55)$$

mit $\alpha^K = (\alpha_1, \dots, \alpha_K)^T$ und $e^K = (e_1, \dots, e_K)^T$. Der Approximationsfehler ist durch $a_T^K = g_T - g_T^K$ gegeben, und für die Repräsentation von X erhalten wir die Regressionsgleichung

$$X = g_T^K(A_T(G)) + a_T^K(A_T(G)) \quad (56)$$

wobei hier der Regressionsfehler gleich dem Approximationsfehler ist.

Für die Berechnung der Koeffizienten ist wieder eine Monte Carlo Simulation nötig. Sei nun $((X_1, A_T(G_1)), \dots, (X_n, A_T(G_n)))$ eine simulierte Stichprobe der Größe n . g_T^K wird wieder durch eine Least Square Regression geschätzt, deren Lösung wie folgt gegeben ist:

$$\hat{g}_T^K = (\hat{\alpha}^K)^T e^K \quad (57)$$

mit

$$\hat{\alpha}^K = ((E_K)^T E_K)^{-1} (E_K)^T \hat{X}$$

wobei $\hat{X} = (X_1, \dots, X_n)^T$ und E_K ist eine $n \times K$ -Matrix, welche in der i -ten Reihe gleich $e^K(A_T(G_i))$ für $i = 1, \dots, n$ stehen hat.

$\hat{\alpha}^K$ entspricht einem gewöhnlichen Least Square Schätzer einer Approximation von X , welche K Basisfunktionen, bewertet zur Zeit T , als unabhängige Variablen verwendet.

Bei diesem Ansatz wird ein Schätzer für die Auszahlungsfunktion durch eine Approximation konstruiert, bei der die Basisfunktionen zum selben Zeitpunkt bewertet werden, wie die Funktion die es zu schätzen gilt. Hat man $X = E_{\mathbb{Q}}[X | \mathcal{F}_T]$ geschätzt, muss der Wert dieser Funktionen zur Zeit t bestimmt werden, um eine Schätzung für die bedingte Erwartungswertfunktion $E_{\mathbb{Q}}[X | \mathcal{F}_t]$ zu bekommen. Da der bedingte Erwartungswert für die Finanzinstrumente, welche als Grundlage für die Basisfunktionen verwendet werden, bekannt ist, kann dieser direkt für den Schätzer von X berechnet werden.

Der RP Schätzer \hat{g}_T^K beinhaltet einen Fehler, welcher sich nur aus dem Approximations-, ist gleich dem Regressions-, und den Monte Carlo Fehler zusammensetzt. Der Monte Carlo Fehler konvergiert für $n \rightarrow \infty$ gegen 0, und auch der Approximationsfehler verschwindet hier für $K \rightarrow \infty$. Im Gegensatz dazu bleibt beim Regressionsfehler der LSMC Methode immer

ein Projektionsfehler, welcher nur 0 ist, wenn man die unabhängigen Variablen der Regression auch zur Zeit T bewertet, was hier nicht der Fall ist, da das Interesse dem diskontierten 1-Jahres Verlust, das heißt dem $E_{\mathbb{Q}}[X|\mathcal{F}_1]$, gilt.

6.2.3 Vor- und Nachteile des RP gegenüber der LSMC

In diesem Unterkapitel werden die Vor- und Nachteile eines replizierenden Portfolios gegen die Least Square Monte Carlo Methode kurz erläutert. Hauptsächlich werden hier die Auswirkungen des Projektionsfehlers, welcher in LSMC einen Teil des Regressionsfehlers ausmacht, betrachtet. Für einen etwas tieferen Einblick in dieses Thema siehe [PS15].

1. Flexibilität des Modells in Bezug auf $t \in [0, T]$:

Bei LSMC wird direkt die bedingte Erwartung zu einem speziellen Zeitpunkt $t^* \leq T$ approximiert, wobei die Basisfunktionen im Gegensatz zur Zielfunktion X zur Zeit t^* bewertet werden. Dies impliziert nicht notwendigerweise eine gute Approximation der bedingten Erwartung durch die selben Basisfunktionen zu einem anderen Zeitpunkt $t \leq T$ mit $t \neq t^*$.

Bei der RP-Methode approximiert man direkt die Auszahlungsfunktion X , auf welche man für ein beliebiges t die bedingte Erwartungsfunktion $E_{\mathbb{Q}}[X|\mathcal{F}_t]$ anwenden kann. Je besser die Replikation der Auszahlungsfunktion zur Zeit T , desto besser wird die bedingte Erwartung zum Zeitpunkt $t < T$ widerspiegelt.

2. Analyse der Varianz:

Beim LSMC-Ansatz wird eine glatte Funktion gesucht, welche die beobachteten Daten am besten repräsentiert. Hier wird versucht eine unbekannt Funktion basierend auf verrauschten Daten zu approximieren. Dies liegt daran, dass die unabhängigen Variablen der Regression zu einem anderen Zeitpunkt als die Abhängige bewertet werden. Die bedingte Varianz des Regressionsfehlers in LSMC ist

$$\text{Var} [a_{0,t}^K(A_t(G)) + p_{0,t}(A_T(G))|\mathcal{F}_t] = \text{Var} [X|\mathcal{F}_t]. \quad (58)$$

Durch die Abhängigkeit vom zugrunde liegenden stochastischen Prozesses und der Struktur von X , kann es sein, dass die bedingte Varianz der zufälligen Auszahlung X zur Zeit T in den unterschiedlichen Beobachtungen zur Zeit t nicht gleich sein muss. Dies kann zu einer Verzerrung des Schätzers führen.

Beim RP-Ansatz wird ein Portfolio gesucht, welches dieselben zeitlichen Zahlungsströme wie die Vorgabe hat. Es wird eine exakte Repräsentation für die Auszahlungsfunktion gesucht, welche auf simulierten Datenpunkten basiert. Diese Regression ist nicht verrauscht, auch wenn der Approximationsfehler nicht 0 ist.

Beim RP-Ansatz ergibt sich für die bedingte Varianz des Approximationsfehlers

$$\text{Var} [a_T^K(A_T(G)) | \mathcal{F}_t] = 0, \quad (59)$$

was zeigt, dass es keine Variationen des Fehlers für unterschiedliche Beobachtung von $A_T(G)$ gibt, daher kann der RP-Ansatz als nicht verzerrt angesehen werden, auch wenn der Approximationsfehler nicht 0 ist.

3. Güte des quadrierten Korrelationskoeffizienten:
Normal ist der quadrierte Korrelationskoeffizient,

$$R^2 = \frac{SSE}{SST} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \in [0, 1],$$

wobei \bar{Y} der empirische Mittelwert der Stichprobe Y_1, \dots, Y_n und $\hat{Y}_1, \dots, \hat{Y}_n$ die Schätzer sind, ein gutes Maß um die Güte der Regression zu messen, durch welchen man den Approximationsfehler analysieren kann. Ein kleines R^2 könnte ein Zeichen für einen fehlenden Risikofaktor sein. Umso größer R^2 ist umso besser ist das Modell.

Für die LSMC-Methode ist der Korrelationskoeffizient jedoch nicht sehr aussagekräftig, was am Projektionsfehler liegt, denn auch wenn der Approximationsfehler gleich 0 ist, gilt

$$R_{lsmc}^2 = 1 - \frac{\sum_{j=1}^{\infty} \alpha_j^2 - \sum_{k=1}^{\infty} \beta_k}{SST} \quad (60)$$

was nur 1 ist, wenn der Projektionsfehler 0 ist, dabei sind die α_j für $j = 1, 2, \dots$ jene Koeffizienten, die man für die perfekte Replikation des Portfolios (also mit Basis $\{e_k\}_{k=1}^{\infty}$), erhält. Die Gleichung (60) ist also die Obergrenze für den Korrelationskoeffizienten in LSMC.

Für die RP-Methode gilt für den Approximationsfehler, welcher nicht gleich 0 ist

$$R_{RP}^2 = 1 - \frac{\sum_{k=K+1}^{\infty} \alpha_k^2}{SST} \quad (61)$$

was für $K \rightarrow \infty$ nach 1 konvergiert. Wenn der Korrelationskoeffizient gleich 1 ist hat man ein perfektes replizierendes Portfolio. Ein kleines R^2 weist hier darauf hin, dass die gewählte Repräsentation nicht gut ist, es könnte sein, dass wichtige Risikofaktoren vernachlässigt worden sind.

4. Asymptotische Unabhängigkeit vom gewählten Maß:

Betrachtet man die Gleichungen für die Berechnung der Koeffizienten (46) und (54), so sieht man, dass die Berechnung von einem Wahrscheinlichkeitsmaß abhängen.

Wählt man im RP eine orthonormale Basis E unter einem zu \mathbb{Q} äquivalenten Wahrscheinlichkeitsmaß $\tilde{\mathbb{Q}}$, erhält man die selben Koeffizienten α zumindest für eine perfekte Basis, das heißt, wenn der Approximationsfehler gleich 0 ist. Ist die Replikation nicht perfekt, dann erhält man bei einem Maßwechsel andere Koeffizienten $\alpha_{\tilde{\mathbb{Q}}}^K$, jedoch konvergieren diese für $K \rightarrow \infty$ gegen die richtigen Koeffizienten. Bei LSMC ändern sich die Koeffizienten durch den Maßwechsel auch wenn der Approximationsfehler gleich 0 ist, da der Projektionsfehler nicht abklingt. Daher ist bei der LSMC-Methode im Gegensatz zur RP Methode keine asymptotische Unabhängigkeit vom Maß gegeben.

5. Pfadabhängigkeit und höher dimensionale Zielfunktionen:

Hat man ein Portfolio zu bewerten, welches pfadabhängig ist, wie es zum Beispiel bei exotischen Optionen der Fall ist, kann das die Dimension des Problems beeinflussen. Je größer die Dimension ist, desto größer ist die Anzahl der Basisfunktionen, um so aufwendiger ist dann natürlich auch die Berechnung.

Bei beiden Methoden gibt es zwei wichtige Schritte um eine gute Repräsentation für die Zielfunktion X mit bedingter Erwartungsfunktion $E_{\mathbb{Q}}[X|\mathcal{F}_t]$ zu finden:

- (a) Identifizieren und zusammenfassen aller Risikofaktoren durch $A_T(G)$ bzw. $A_t(G)$, welche die Zielfunktion beeinflussen.
- (b) Finden einer Basis, mit welcher $A_t(G)$ für LSMC bzw. $A_T(G)$ für RP dargestellt werden kann.

Die Komplexität des Problems hängt von der Anzahl und dem Typus der Risikofaktoren ab, von welchen die Zielfunktion abhängt.

Für die RP-Methode können verschiedene $A_T(G)$ mit verschiedenen Dimensionen gewählt werden. Rechnerisch wäre natürlich die kleinere Dimension der größeren vorzuziehen. Es gibt dabei jedoch einen wesentlichen Nachteil, da die Berechnung der bedingten Erwartungsfunktion bei einer kleineren Dimension komplizierter sein kann. Da es gerade bei RP wichtig ist diese einfach zu halten, sollte man die größere Dimension bevorzugen.

Da bei der LSMC-Methode die Basis direkt zum Zeitpunkt t berechnet wird, braucht man darauf nicht mehr den bedingten Erwartungswertoperator anwenden, daher kann hier immer die kleinere Dimension für $A_t(G)$ verwendet werden.

Wie nun aufgezeigt wurde, hat das replizierende Portfolio einige Vorteile gegenüber der Least Square Monte Carlo Methode, da der Regressionsfehler, durch das Fehlen des Projektionsfehlers, bei Annäherung an eine perfekte Basis, bessere Eigenschaften aufweist, und dadurch einfacher zu analysieren ist. Auch wenn die Pfadabhängigkeit das Problem ein replizierendes Portfolio zu finden etwas aufwendiger macht, lohnt es sich, da die anderen Vorteile überwiegen. Im folgenden wird die Methode des replizierenden Portfolios bei einer festen Anzahl der Finanzinstrumente, welche für die Replikation zur Verfügung stehen, genauer untersucht.

7 Replizierendes Portfolio

Dieses Kapitel behandelt die Theorie des dynamischen replizierenden Portfolios, wie sie im Paper [CF17] erläutert wird.

Da die Anzahl der Finanzinstrumente, welche am Markt zur Verfügung stehen und effektiv berechnet werden können, beschränkt ist, wird nun angenommen, dass zusätzlich zum Numéraire m Finanzinstrumente mit diskontiertem Gewinn-Prozess

$$G_t = (G_{1t}, \dots, G_{mt})^T$$

zum Zeitpunkt t , für die Berechnung eines replizierenden Portfolios verwendet werden können. Da Arbitragefreiheit gewünscht wird, benötigt man, dass G_t ein \mathbb{Q} -Martingal ist. Man versucht X und dadurch L durch ein Portfolio zu approximieren, welches aus diesen m Finanzinstrumenten besteht.

Es wird nun in erster Linie der Fall behandelt, wo es möglich ist das replizierende Portfolio zu gewissen Zeitpunkten auszubalancieren, und auch eine Umschichtung zwischen den zur Verfügung stehenden Finanzinstrumenten soll ermöglicht werden.

Die Theorie in diesem Kapitel wird sowohl für das reale Maß als auch das risikoneutrale Maß besprochen. Daher sei \mathbb{M} gleich \mathbb{P} oder \mathbb{Q} .

7.1 Dynamisches replizierendes Portfolio

Beim dynamischen replizierenden Portfolio wird eine Partition des Zeithorizontes $[0, T]$, $0 = t_0 < t_1 < \dots < t_N = T$, fixiert, wobei $t_j = 1$ für ein j gilt. Des weiteren wird

$$\Delta G_{ij} = G_{it_j} - G_{it_{j-1}}$$

für den Gewinn des Instrumentes i im Intervall $(t_{j-1}, t_j]$ geschrieben. Die Portfoliostrategien sind multilinear in den laufenden Gewinnen ΔG_{ij} , was heißt, dass sie linear sind in den Produkten dieser Gewinne, was in der stochastischen Analysis auch als Chaosentwicklung bekannt ist. Formal kann das wie folgt definiert werden:

Es sei \mathcal{P} eine Familie von Paaren $(\mathcal{I}, \mathcal{J})$, wobei $\mathcal{J} \subset \{1, \dots, N\}$ ist und $\mathcal{I} : \mathcal{J} \rightarrow \{1, \dots, m\}$ eine Abbildung. Für $(\mathcal{I}, \mathcal{J}) \in \mathcal{P}$ wird das dazugehörige Produkt der Gewinne wie folgt definiert

$$\Delta G_{(\mathcal{I}, \mathcal{J})} = \prod_{j \in \mathcal{J}} \Delta G_{\mathcal{I}(j), j},$$

und es wird angenommen, dass $\Delta G_{(\mathcal{I}, \mathcal{J})} \in L^2(\mathbb{M})$ ist.

Da G_t ein \mathbb{Q} -Martingal ist gilt

$$E_{\mathbb{Q}} [\Delta G_{(\mathcal{I}, \mathcal{J})} | \mathcal{F}_{t_j}] = 0, \text{ für alle } j < \min \mathcal{J}. \quad (62)$$

Es sei nun $\Delta G_{(\mathcal{I}, \mathcal{J})}$, die Basis aus derer das replizierende Portfolio berechnet werden soll. Um diese Definition der Basis etwas anschaulicher zu erklären, betrachte man folgendes Beispiel:

Beispiel 7.1. *Es sei $0 = t_0 < t_1 < t_2 = T$, und es stehen auf dem Finanzmarkt zwei Finanzinstrumente zur Verfügung, das heißt $G = (G_1, G_2)^T$. Es gilt nun $\mathcal{J} \subset \{1, 2\}$ und $\mathcal{I} : \mathcal{J} \rightarrow \{1, 2\}$.*

Die Basis für $|\mathcal{J}| = 1$, wäre dann

$$\begin{pmatrix} G_{1t_1} - G_{1t_0} \\ G_{1t_2} - G_{1t_1} \\ G_{2t_1} - G_{2t_0} \\ G_{2t_2} - G_{2t_1} \end{pmatrix}.$$

Für $|\mathcal{J}| = 2$ würde man die Basis

$$\begin{pmatrix} (G_{1t_1} - G_{1t_0})(G_{1t_2} - G_{1t_1}) \\ (G_{1t_1} - G_{1t_0})(G_{2t_2} - G_{2t_1}) \\ (G_{2t_1} - G_{2t_0})(G_{1t_2} - G_{1t_1}) \\ (G_{2t_1} - G_{2t_0})(G_{2t_2} - G_{2t_1}) \end{pmatrix}$$

bekommen. Es ist auch möglich diese Basen miteinander zu kombinieren, dann hat man jene Basis, welche alle möglichen Basiselemente beinhaltet. Verbessern kann man diese Basis nur indem man weitere Finanzinstrumente hinzufügt, oder die Anzahl der t_i erhöht.

Es gilt, dass man für eine beliebige Wahl deterministischer Koeffizienten $\alpha = \{\alpha_{(\mathcal{I}, \mathcal{J})} | (\mathcal{I}, \mathcal{J}) \in \mathcal{P}\} \in \mathbb{R}^{|\mathcal{P}|}$ und einen Anfangswert v ein selbstfinanzierendes Portfolio mit Werteprozess

$$V_t^{v, \alpha} = v + \sum_{(\mathcal{I}, \mathcal{J}) \in \mathcal{P} | t_{\max \mathcal{J}} \leq t} \alpha_{(\mathcal{I}, \mathcal{J})} \Delta G_{(\mathcal{I}, \mathcal{J})}$$

erhält. Mit (62) folgt für das selbstfinanzierende Portfolio $V_t^{v, \alpha}$, dass es ein \mathbb{Q} -Martingal ist.

Damit ist der Grundbaustein für die dynamische Portfolio Replikation gelegt.

Um die Notation im folgenden zu Vereinfachen setzt man

$$A = (\Delta G_{(\mathcal{I}, \mathcal{J})} | t_{\max \mathcal{J}} \leq 1)^T, \quad B = (\Delta G_{(\mathcal{I}, \mathcal{J})} | t_{\max \mathcal{J}} > 1)^T$$

für den Zufallsvektor aller möglichen Portfoliogewinne bis zu einem Jahr und nach einem Jahr. Auf den gleichen Weg wird der Koeffizienten Vektor des Portfolios $\alpha^T = (\alpha_A^T, \alpha_B^T)$ zerlegt. Damit erleichtert man sich die Berechnung der benötigten Werte des Portfolios $V_t^{v, \alpha}$ zu den Zeitpunkten $t = 0, 1, T$. Man erhält sie ganz einfach durch

$$V_0^{v,\alpha} = v, \quad V_1^{v,\alpha} = v + \alpha_A^T A, \quad V_T^{v,\alpha} = v + \alpha_A^T A + \alpha_B^T B.$$

Für die Berechnung der Koeffizienten muss das $L^2(\mathbb{M})$ -Minimierungsproblem

$$\min_{(v,\alpha) \in \mathbb{R}^{|\mathcal{P}|+1}} \|Z - V_T^{v,\alpha}\|_{L^2(\mathbb{M})}. \quad (63)$$

gelöst werden. Das daraus resultierende selbstfinanzierende Portfolio $V_t^{v,\alpha}$ wird im folgenden als dynamisches replizierendes Portfolio bezeichnet. Im Allgemeinen ist die Replikation für Z jedoch nicht perfekt.

Bemerkung 7.1. *Die Replikation ist perfekt, wenn die Menge $\{\Delta G_{(\mathcal{I},\mathcal{J})} \mid (\mathcal{I},\mathcal{J}) \in \mathcal{P}\}$ eine Basis des $L^2(\mathbb{Q})$ bildet. Diese Eigenschaft erfüllen nicht alle Martingale. Ein Martingal, welches diese erfüllt, ist die Brownsche Bewegung, wenn man die Anzahl der t_j gegen unendlich gehen lässt.*

Die Lösung des Minimierungsproblems ist der Endwert von $V_t^{v,\alpha}$ und kann als $L^2(\mathbb{M})$ -Projektion von Z auf $\{1, A, B\}$ interpretiert werden. Man kann sie für das risikoneutrale Maß \mathbb{Q} und das reale Maß \mathbb{P} wie folgt berechnen.

- Für $\mathbb{M} = \mathbb{P}$ ist die formale Lösung von (63) im Allgemeinen durch den Ausdruck

$$\begin{pmatrix} v^{\mathbb{P}} \\ \alpha_A^{\mathbb{P}} \\ \alpha_B^{\mathbb{P}} \end{pmatrix} = \mathcal{M}^{-1} E_{\mathbb{P}} \left[\begin{pmatrix} Z \\ AZ \\ BZ \end{pmatrix} \right] \quad (64)$$

mit Gramscher Matrix

$$\mathcal{M} = E_{\mathbb{P}} \left[\begin{pmatrix} 1 & A^T & B^T \\ A & AA^T & AB^T \\ B & BA^T & BB^T \end{pmatrix} \right].$$

gegeben.

- Im Fall $\mathbb{M} = \mathbb{Q}$ vereinfacht sich die Darstellung von oben, da wegen der Martingaleigenschaft (62) $E_{\mathbb{Q}}[A] = E_{\mathbb{Q}}[B] = E_{\mathbb{Q}}[AB^T] = 0$ gilt. Da von einer fairen Bepreisung unter dem Martingalmaß ausgegangen wird, gilt auch $E_{\mathbb{Q}}[Z] = 0$. Die Lösung des Minimierungsproblems ist daher durch

$$v^{\mathbb{Q}} = 0, \quad \begin{pmatrix} \alpha_A^{\mathbb{Q}} \\ \alpha_B^{\mathbb{Q}} \end{pmatrix} = \mathcal{N}^{-1} E_{\mathbb{Q}} \left[\begin{pmatrix} AZ \\ BZ \end{pmatrix} \right] \quad (65)$$

gegeben, wobei hier die Gramsche Matrix eine Blockdiagonalmatrix ist

$$\mathcal{N} = E_{\mathbb{Q}} \left[\begin{pmatrix} AA^T & 0 \\ 0 & BB^T \end{pmatrix} \right]$$

Tatsächlich erhält man in diesem Fall $\alpha_A^{\mathbb{Q}}$ bzw. $\alpha_B^{\mathbb{Q}}$ jeweils auch durch direkte Projektion von Z auf A oder B . In diesem Fall kann man zwei unabhängige Optimierungsprobleme betrachten.

In der Praxis könnte die Berechnung der Inversen der Gramschen Matrizen \mathcal{N} und \mathcal{M} für große Dimensionen, Probleme machen, da sie aufgrund der möglichen starken Korrelation zwischen den Finanzinstrumenten G_t , nahezu singular sein können. Finanzinstrumente, für welche die Matrizen explizit gegeben sind, sind daher zu bevorzugen.

Sollten die Matrizen nicht explizit zu berechnen sein, sollte man dies bei der Wahl der Anzahl der Basisfunktionen bedenken. Genauer wird dies anhand der Beispiele in Kapitel 9 analysiert.

7.2 Statisches Replizierendes Portfolio

Als statisches Replizierendes Portfolio wird jenes Portfolio bezeichnet, welches nur zum Zeitpunkt T betrachtet wird, dass heißt es ist nicht möglich, das Portfolio umzuschichten, und die Pfadabhängigkeit wird auch nicht berücksichtigt. Betrachtet man nun diese Methode mit den fixen Basisinstrumenten

$$G_t = (G_{1t}, \dots, G_{mt}),$$

wie es in einem Beispiel des Papers [CF17] gemacht wurde, dann gilt, dass $N = 1$ ist, und das Portfolio wird über den Zeithorizont $[0, T]$ nicht verändert. Man setzt $A = G_1 - G_0$, $B = G_T - G_1$, dann ist $A + B = G_T - G_0$, und es ist $\alpha_A = \alpha_B$. In diesem Fall wird die Lösung des $L^2(\mathbb{M})$ -Minimierungsproblems

$$\min_{(v, \alpha) \in \mathbb{R}^{1+m}} \|Z - v - \alpha^T(A + B)\|_{L^2(\mathbb{M})}.$$

gesucht. Die Lösung dieses Problems wird wieder separat für die zwei Wahrscheinlichkeitsmaße betrachtet.

- $\mathbb{M} = \mathbb{P}$

$$\begin{pmatrix} v^{\mathbb{P}} \\ \alpha^{\mathbb{P}} \end{pmatrix} = \mathcal{M}^{-1} E_{\mathbb{P}} \left[\begin{pmatrix} Z \\ (A + B) Z \end{pmatrix} \right] \quad (66)$$

mit Gram-Matrix

$$\mathcal{M} = E_{\mathbb{P}} \left[\begin{pmatrix} 1 & (A + B)^T \\ (A + B) & (A + B)(A + B)^T \end{pmatrix} \right].$$

- $\mathbb{M} = \mathbb{Q}$

$$v^{\mathbb{Q}} = 0, \quad \alpha^{\mathbb{Q}} = \mathcal{N}^{-1} E_{\mathbb{Q}} [(A + B) Z] \quad (67)$$

wobei hier die reduzierte Gram-Matrix wie folgt gegeben ist:

$$\mathcal{N} = E_{\mathbb{Q}} \left[(A + B)(A + B)^T \right].$$

7.3 Analyse des Approximationsfehler

Es wird nun der Approximationsfehler analysiert, indem versucht wird eine gute Schranke zu finden, wenn man eine feste Anzahl von Finanzinstrumenten zur Verfügung hat. Der Fehler ist durch das Residuum der $L^2(\mathbb{M})$ -Projektion

$$\epsilon^{\mathbb{M}} = Z - v^{\mathbb{M}} - (\alpha_A^{\mathbb{M}})^T A - (\alpha_B^{\mathbb{M}})^T B$$

gegeben. Damit kann man den Einjahresverlust L durch

$$L = E_{\mathbb{Q}}[Z|\mathcal{F}_1] = v^{\mathbb{M}} + (\alpha_A^{\mathbb{M}})^T A + E_{\mathbb{Q}}[\epsilon^{\mathbb{M}}|\mathcal{F}_1],$$

ausdrücken, wobei wegen der Martingaleigenschaft $E_{\mathbb{Q}}[(\alpha_B^{\mathbb{B}})^T B|\mathcal{F}_1] = 0$ gilt.

Da der $E_{\mathbb{Q}}[\epsilon^{\mathbb{M}}|\mathcal{F}_1]$ auch nicht direkt berechenbar ist, bieten sich folgende zwei Darstellungen als Schätzer für L an.

$$L_1^{\mathbb{M}} = v^{\mathbb{M}} + (\alpha_A^{\mathbb{M}})^T A \tag{68}$$

$$L_2^{\mathbb{M}} = v^{\mathbb{M}} + (\alpha_A^{\mathbb{M}})^T A + \epsilon^{\mathbb{M}} = Z - (\alpha_B^{\mathbb{M}})^T B \tag{69}$$

Das erforderte Kapital wird über die Schätzer berechnet und ist durch

$$\begin{aligned} K_1^{\mathbb{M}} &= \rho(L_1^{\mathbb{M}}) = v^{\mathbb{M}} + \rho((\alpha_A^{\mathbb{M}})^T A) \\ K_2^{\mathbb{M}} &= \rho(L_2^{\mathbb{M}}) = \rho(Z - (\alpha_B^{\mathbb{M}})^T B) \end{aligned}$$

gegeben, wobei ρ ein Risikomaß ist, also bei diesen Betrachtungen der Expected Shortfall oder der Value at Risk. Der Schätzer $L_2^{\mathbb{M}}$ ist nur dann ein guter Schätzer für L , wenn eine der äquivalenten Bedingungen des folgenden Lemmas erfüllt sind.

Lemma 7.1. *Folgende Aussagen sind äquivalent:*

1. $E_{\mathbb{Q}}[\epsilon^{\mathbb{M}}|\mathcal{F}_1] = \epsilon^{\mathbb{M}}$;
2. $\epsilon^{\mathbb{M}}$ ist \mathcal{F}_1 -messbar;
3. Die Cash-Flows nach $t = 1$ werden von den Instrumenten G_t aufgespannt

$$Z - L = (\alpha_B^{\mathbb{M}})^T B.$$

Beweis. Die Äquivalenz der ersten beiden Aussagen ist eine Eigenschaft der bedingten Erwartung.

Es wird nun 1. \Leftrightarrow 3. gezeigt, was ganz einfach durch einsetzen folgt:

1. \Rightarrow 3.:

$$Z - L = Z - v^{\mathbb{M}} - (\alpha_A^{\mathbb{M}})^T A - E_{\mathbb{Q}}[\epsilon^{\mathbb{M}}|\mathcal{F}_1]$$

$$\begin{aligned}
&= Z - v^{\mathbb{M}} - (\alpha_A^{\mathbb{M}})^T A - \epsilon^{\mathbb{M}} \\
&= Z - v^{\mathbb{M}} - (\alpha_A^{\mathbb{M}})^T A - Z + (\alpha_A^{\mathbb{M}})^T A + (\alpha_B^{\mathbb{M}})^T B = (\alpha_B^{\mathbb{M}})^T B
\end{aligned}$$

3. \Rightarrow 1.

$$\begin{aligned}
E_{\mathbb{Q}} \left[\epsilon^{\mathbb{M}} | \mathcal{F}_1 \right] &= L - v^{\mathbb{M}} - (\alpha_A^{\mathbb{M}})^T A + Z - Z \\
&= Z - v^{\mathbb{M}} - (\alpha_A^{\mathbb{M}})^T A - \underbrace{(Z - L)}_{(\alpha_B^{\mathbb{M}})^T B} = \epsilon^{\mathbb{M}}
\end{aligned}$$

□

Im Kapitel 9 wird nur der Schätzer $L_1^{\mathbb{M}}$ betrachtet. Jedoch wird bei der theoretischen Analyse der Fehler, wegen der Vollständigkeit weiterhin auf beide Varianten eingegangen.

Aus dem folgenden Satz erhält man eine obere Schranke für den $L^1(\mathbb{P})$ -Approximationsfehler beider Stellvertreter.

Satz 7.1. *Die $L^1(\mathbb{P})$ -Approximationsfehler sind durch*

$$\left\| L - L_1^{\mathbb{M}} \right\|_{L^1(\mathbb{P})} \leq C_2^{\mathbb{M}} \left\| \epsilon^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} \quad (70)$$

und

$$\begin{aligned}
\left\| L - L_2^{\mathbb{M}} \right\|_{L^1(\mathbb{P})} &\leq \left\| E_{\mathbb{Q}} \left[\epsilon^{\mathbb{M}} | \mathcal{F}_1 \right] - \epsilon^{\mathbb{M}} \right\|_{L^1(\mathbb{P})} \\
&\leq \begin{cases} (1 + C_2^{\mathbb{P}}) \left\| \epsilon^{\mathbb{M}} \right\|_{L^2(\mathbb{P})} \\ C_1^{\mathbb{Q}} \left\| E_{\mathbb{Q}} \left[\epsilon^{\mathbb{M}} | \mathcal{F}_1 \right] - \epsilon^{\mathbb{M}} \right\|_{L^2(\mathbb{Q})} \leq C_1^{\mathbb{Q}} \left\| \epsilon^{\mathbb{M}} \right\|_{L^2(\mathbb{Q})} \end{cases} \quad (71)
\end{aligned}$$

Beweis. Im Beweis wird das Lemma 2.4 verwendet.

$$\begin{aligned}
\left\| L - L_1^{\mathbb{M}} \right\|_{L^1(\mathbb{P})} &= \left\| E_{\mathbb{Q}} \left[\epsilon^{\mathbb{M}} | \mathcal{F}_1 \right] \right\|_{L^1(\mathbb{P})} \\
&\stackrel{\text{Lemma 2.4}}{\leq} C_2^{\mathbb{M}} \left\| \epsilon^{\mathbb{M}} \right\|_{L^2(\mathbb{M})}
\end{aligned}$$

Die zweite Ungleichung bekommt man durch folgende Ungleichungskette:

$$\begin{aligned}
\left\| L - L_2^{\mathbb{M}} \right\|_{L^1(\mathbb{P})} &\leq \left\| E_{\mathbb{Q}} \left[\epsilon^{\mathbb{M}} | \mathcal{F}_1 \right] - \epsilon^{\mathbb{M}} \right\|_{L^1(\mathbb{P})} \\
&\leq \left\| E_{\mathbb{Q}} \left[\epsilon^{\mathbb{M}} | \mathcal{F}_1 \right] \right\|_{L^1(\mathbb{P})} + \left\| \epsilon^{\mathbb{M}} \right\|_{L^1(\mathbb{P})}
\end{aligned}$$

und der Hölder Ungleichung. □

Aus dem Satz folgt, dass die replizierende Portfolio Methode, basierend auf $L_i^{\mathbb{M}}$ für $i = 1, 2$, $\mathbb{M} = \mathbb{P}$ oder \mathbb{Q} und $\rho = VaR_\alpha$ oder ES_α , asymptotisch konsistent ist, wenn die Anzahl $|\mathcal{P}|$ erhöht wird. Was heißt, dass dann $\{1, A, B\}$ eine Basis des $L^2(\mathbb{M})$ bildet.

Mit (17) und (19) impliziert dies, dass der Approximationsfehler des Kapitals gegen Null konvergiert, also

$$K_i^{\mathbb{M}} \rightarrow K,$$

wobei für den Value at Risk, $\rho = VaR_\alpha$, angenommen werden muss, dass $q_\alpha^-(L) = q_\alpha^+(L)$. Dabei ist zu beachten, dass das Anwachsen der Anzahl von \mathcal{P} sowohl durch vergrößern der Anzahl der Finanzinstrumente G_t erreicht werden kann, als auch durch erhöhen der Anzahl der Zeitschritte N und/oder dem Grad der Pfadabhängigkeit $|\mathcal{J}|$.

Für den Expected Shortfall, $\rho = ES_\alpha$, könnte man (70) und (71) mit der Lipschitz Bedingung (19) kombinieren, um eine obere Schranke für den Approximationsfehler des Eigenkapitalbedarfs $|K - K_i^{\mathbb{M}}|$ zu erhalten.

Korollar 7.1. *Für den Expected Shortfall, ES_α erhält man*

$$|K - K_1^{\mathbb{M}}| \leq \frac{1}{1 - \alpha} \left(C_2^{\mathbb{M}} \|\epsilon^{\mathbb{M}}\|_{L^2(\mathbb{M})} \right)$$

und

$$|K - K_2^{\mathbb{M}}| \leq \begin{cases} \frac{1}{1 - \alpha} (1 + C_2^{\mathbb{P}}) \|\epsilon^{\mathbb{M}}\|_{L^2(\mathbb{P})} \\ \frac{1}{1 - \alpha} \left(C_1^{\mathbb{Q}} \|\epsilon^{\mathbb{M}}\|_{L^2(\mathbb{Q})} \right) \end{cases}$$

eine obere Schranke für den Approximationsfehler des Kapitals.

Obwohl diese Grenzen asymptotisch stark sind, sind sie nicht scharf.

Eine äquivalente Aussage zu Korollar 7.1 kann für den Value at Risk angegeben werden, welche direkt aus der Proposition 4.1 und dem Satz 7.1 folgt.

Korollar 7.2. *Für den Einjahresgewinn L und $\alpha \in (0, 1)$ seien die Voraussetzungen von Lemma 4.2 für ein $\gamma > 0$ erfüllt. Dann existiert $\delta > 0$, so dass unter der Bedingung $d_p(L, L_i^{\mathbb{M}}) < \delta$ für $i = 1, 2$*

$$|K - K_1^{\mathbb{M}}| \leq \left(1 + \frac{1}{\gamma} \right) \left(C_2^{\mathbb{M}} \|\epsilon^{\mathbb{M}}\|_{L_2(\mathbb{M})} \right)^{\frac{1}{2}}$$

und

$$|K - K_2^{\mathbb{M}}| \leq \begin{cases} \left(1 + \frac{1}{\gamma} \right) \left((C_1^{\mathbb{P}} + C_2^{\mathbb{P}}) \|\epsilon^{\mathbb{M}}\|_{L_2(\mathbb{P})} \right)^{\frac{1}{2}} \\ \left(1 + \frac{1}{\gamma} \right) \left(C_1^{\mathbb{Q}} \|\epsilon^{\mathbb{M}}\|_{L_2(\mathbb{Q})} \right)^{\frac{1}{2}} \end{cases}$$

gilt.

8 MC-Methoden angewendet auf RP

Die Lösung des $L^2(\mathbb{M})$ -Optimierungsproblem (63) hängt von Zufallsvariablen ab, deren Verteilung nicht bekannt ist, daher ist es nötig empirische Regression anzuwenden, was bedeutet, dass das Problem mit Monte Carlo Simulation gelöst werden muss. Dazu wird angenommen, dass $AZ, BZ \in L^2(\mathbb{M})$. Man simuliere n i.i.d. Kopien

$$(A^{(j)}, B^{(j)}, Z^{(j)}), \quad j = 1, \dots, n \quad (72)$$

von (A, B, Z) unter \mathbb{M} . Mit \mathcal{G} wird die σ -Algebra bezeichnet, welche durch diese Stichprobe erzeugt wird.

Die Lösungen für \mathbb{P} und \mathbb{Q} sind wie folgt gegeben:

- Für $\mathbb{M} = \mathbb{P}$ erhält man nun bezüglich (64) die unverzerrten Schätzer

$$\begin{pmatrix} \hat{v}^{\mathbb{P}} \\ \hat{\alpha}_A^{\mathbb{P}} \\ \hat{\alpha}_B^{\mathbb{P}} \end{pmatrix} = \mathcal{M}^{-1} \frac{1}{n} \sum_{j=1}^n \begin{pmatrix} Z^{(j)} A^{(j)} Z^{(j)} \\ B^{(j)} Z^{(j)} \end{pmatrix}$$

für $(v^{\mathbb{P}}, \alpha_A^{\mathbb{P}}, \alpha_B^{\mathbb{P}})$. Aus dem Gesetz der großen Zahlen folgt, dass

$$(\hat{v}^{\mathbb{P}}, \hat{\alpha}_A^{\mathbb{P}}, \hat{\alpha}_B^{\mathbb{P}}) \rightarrow (v^{\mathbb{P}}, \alpha_A^{\mathbb{P}}, \alpha_B^{\mathbb{P}}) \quad f.s., \quad \text{für } n \rightarrow \infty. \quad (73)$$

Der zentrale Grenzwertsatz impliziert, dass $(\hat{v}^{\mathbb{P}}, \hat{\alpha}_A^{\mathbb{P}}, \hat{\alpha}_B^{\mathbb{P}})$, asymptotisch für große n , normalverteilt ist mit Erwartungswert $(v^{\mathbb{P}}, \alpha_A^{\mathbb{P}}, \alpha_B^{\mathbb{P}})$ und Kovarianzmatrix

$$Cov_{\mathbb{P}} \left[\begin{pmatrix} \hat{v}^{\mathbb{P}} \\ \hat{\alpha}_A^{\mathbb{P}} \\ \hat{\alpha}_B^{\mathbb{P}} \end{pmatrix} \right] = \frac{1}{n} \mathcal{M}^{-1} Cov_{\mathbb{P}} \left[\begin{pmatrix} Z \\ AZ \\ BZ \end{pmatrix} \right] \mathcal{M}^{-1} = \frac{1}{n} C^{\mathbb{P}} \quad (74)$$

- Für $\mathbb{M} = \mathbb{Q}$ vereinfacht sich der obige Ausdruck wieder etwas. In Bezug auf (65) bekommt man den unverzerrten Schätzer

$$\hat{v}^{\mathbb{Q}} = 0, \quad \begin{pmatrix} \hat{\alpha}_A^{\mathbb{Q}} \\ \hat{\alpha}_B^{\mathbb{Q}} \end{pmatrix} = \mathcal{N}^{-1} \frac{1}{n} \sum_{j=1}^n \begin{pmatrix} A^{(j)} Z^{(j)} \\ B^{(j)} Z^{(j)} \end{pmatrix} \quad (75)$$

für die Portfoliostrategie $v^{\mathbb{Q}} = 0$ und $(\alpha_A^{\mathbb{Q}}, \alpha_B^{\mathbb{Q}})$. Aus dem Gesetz der Großen Zahlen folgt auch hier

$$(\hat{\alpha}_A^{\mathbb{Q}}, \hat{\alpha}_B^{\mathbb{Q}}) \rightarrow (\alpha_A^{\mathbb{Q}}, \alpha_B^{\mathbb{Q}}), \quad f.s. \quad \text{für } n \rightarrow \infty. \quad (76)$$

Aus dem zentralen Grenzwertsatz folgt wie im Fall $\mathbb{M} = \mathbb{P}$, dass für große n der Vektor $(\hat{\alpha}_A^{\mathbb{Q}}, \hat{\alpha}_B^{\mathbb{Q}})$ asymptotisch normalverteilt ist mit Erwartungswert $(\alpha_A^{\mathbb{Q}}, \alpha_B^{\mathbb{Q}})$ und Kovarianz-Matrix

$$Cov_{\mathbb{Q}} \left[\begin{pmatrix} \hat{\alpha}_A^{\mathbb{Q}} \\ \hat{\alpha}_B^{\mathbb{Q}} \end{pmatrix} \right] = \frac{1}{n} \mathcal{N}^{-1} Cov_{\mathbb{Q}} \left[\begin{pmatrix} AZ \\ BZ \end{pmatrix} \right] \mathcal{N}^{-1} = \frac{1}{n} C^{\mathbb{Q}} \quad (77)$$

Damit bekommt man die Schätzer für die Stellvertreter des Eigenkapitalbedarfs $K_i^{\mathbb{M}}$ für $i = 1, 2$

$$\begin{aligned} \hat{K}_1^{\mathbb{M}} &= \hat{v}^{\mathbb{M}} + \rho \left((\hat{\alpha}_A^{\mathbb{M}})^T A | \mathcal{G} \right) \\ \hat{K}_2^{\mathbb{M}} &= \rho \left(Z - (\hat{\alpha}_B^{\mathbb{M}})^T B | \mathcal{G} \right), \end{aligned} \quad (78)$$

wobei $\rho(\cdot | \mathcal{G})$ bezeichnet den bedingten Value at Risk oder Expected Shortfall gegeben die σ -Algebra \mathcal{G} , welche durch (72) erzeugt wird. Diese Schätzer sind stark konsistente Schätzer.

Satz 8.1. *Es gilt $\hat{K}_i^{\mathbb{M}} \rightarrow K_i^{\mathbb{M}}$ f.s. für $n \rightarrow \infty$ und $i = 1, 2$, wobei für den Value at Risk, $\rho = VaR_{\alpha}$, angenommen wird, dass $q_{L_i^{\mathbb{M}}}^{-}(1 - \alpha) = q_{L_i^{\mathbb{M}}}^{+}(1 - \alpha)$.*

Beweis. Der Satz folgt aus (17),(19) kombiniert mit (73),(76). \square

Der totale Schätzfehler des Eigenkapitalbedarf setzt sich wie folgt zusammen:

$$\left\| K - \hat{K}_i^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} \leq \left| K - K_i^{\mathbb{M}} \right| + \left\| K_i^{\mathbb{M}} - \hat{K}_i^{\mathbb{M}} \right\|_{L^2(\mathbb{M})}, \quad (79)$$

hierbei ist der erste Term auf der rechten Seite der Approximationsfehler, für welchen sowohl für den Expected Shortfall im Korollar 7.1 als auch für den Value at Risk mit Korollar 7.2 eine obere Schranke hergeleitet wurde. Der zweite Term auf der rechten Seite ist der Monte Carlo Fehler, für welchen folgende obere Schranke gilt.

Satz 8.2. *Für den Expected Shortfall, $\rho = ES_{\alpha}$ ist der Monte Carlo Fehler asymptotisch für große n durch*

$$\left\| K_i^{\mathbb{M}} - \hat{K}_i^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} \leq \sqrt{\frac{1}{n}} MCE_{i,ES}^{\mathbb{M}}$$

mit Konstanten $MCE_{i,ES}^{\mathbb{M}}$ beschränkt welche wie folgt gegeben sind:

$$MCE_{i,ES}^{\mathbb{M}} = \begin{cases} \sqrt{C_{vv}^{\mathbb{P}} + \frac{1}{1-\alpha} \text{tr} (C_{AA}^{\mathbb{P}} E_{\mathbb{P}} [AA^T])}, & \text{für } i = 1 \text{ und } \mathbb{M} = \mathbb{P}, \\ \frac{1}{1-\alpha} \sqrt{\text{tr} (C_{AA}^{\mathbb{Q}} E_{\mathbb{P}} [AA^T])}, & \text{für } i = 1 \text{ und } \mathbb{M} = \mathbb{Q}, \\ \frac{1}{1-\alpha} \sqrt{\text{tr} (C_{BB}^{\mathbb{M}} E_{\mathbb{P}} [BB^T])}, & \text{für } i = 2, \end{cases}$$

wobei $C_{vv}^{\mathbb{M}}$, $C_{AA}^{\mathbb{M}}$ und $C_{BB}^{\mathbb{M}}$ die diagonalen 1×1 -, $|A| \times |A|$ - und $|B| \times |B|$ -Blöcke der Matrix $C^{\mathbb{M}}$ sind.

Beweis. Für $i = 1$ gilt

$$\begin{aligned} \left\| K_i^{\mathbb{M}} - \hat{K}_i^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} &= \left\| v^{\mathbb{M}} - \hat{v}^{\mathbb{M}} + \rho \left((\alpha_A^{\mathbb{M}})^T A \right) - \rho \left((\hat{\alpha}_A^{\mathbb{M}})^T A | \mathcal{G} \right) \right\|_{L^2(\mathbb{M})} \\ &\leq \left\| v^{\mathbb{M}} - \hat{v}^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} + \left\| \rho \left((\alpha_A^{\mathbb{M}})^T A \right) - \rho \left((\hat{\alpha}_A^{\mathbb{M}})^T A | \mathcal{G} \right) \right\|_{L^2(\mathbb{M})} = (*). \end{aligned}$$

Für großes n ist $v^{\mathbb{M}} - \hat{v}^{\mathbb{M}}$ annähernd normalverteilt mit Erwartungswert 0 und Varianz $\frac{1}{n} C_{vv}^{\mathbb{M}}$. Damit, mit der Lipschitzbedingung des Expected Shortfalls (19) und da $f(x) = x^2$ konvex ist, gilt:

$$(*) \leq \sqrt{\frac{1}{n} \sqrt{C_{vv}^{\mathbb{M}}} + \frac{1}{1+\alpha} \sqrt{E_{\mathbb{M}} \left[E_{\mathbb{P}} \left[((\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}})^T A)^2 | \mathcal{G} \right] \right]}}$$

$$\stackrel{\substack{(\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}}) \text{ } \mathcal{G}\text{-m.b.} \\ AA^T \text{ u.a. von } \mathcal{G}}}{AA^T} \sqrt{\frac{1}{n} \sqrt{C_{vv}^{\mathbb{M}}} + \frac{1}{1+\alpha} \sqrt{E_{\mathbb{M}} \left[(\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}})^T E_{\mathbb{P}} [AA^T] (\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}}) \right]}} = (+).$$

Da für zwei Vektoren x, y

$$x^T y = \text{tr} (x^T y) = \text{tr} (y x^T) = \text{tr} (x y^T)$$

gilt, wobei $\text{tr}(\cdot)$ als die Spur einer Matrix definiert ist, folgt daraus

$$(+) = \sqrt{\frac{1}{n} \sqrt{C_{vv}^{\mathbb{M}}} + \frac{1}{1+\alpha} \sqrt{E_{\mathbb{M}} \left[\text{tr} (E_{\mathbb{P}} [AA^T]) (\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}}) (\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}})^T \right]}}$$

Woraus man nun mit dem Zentralen Grenzwertsatz für $\mathbb{M} = \mathbb{P}$ bzw. $\mathbb{M} = \mathbb{Q}$ das gewünschte Ergebnis erhält.

Die Ungleichung für $i = 2$ folgt analog. □

Für den Value at Risk benötigt man wieder die zusätzlichen Voraussetzungen des Lemmas 4.2.

Satz 8.3. Für $L_i^{\mathbb{M}}$ seien die Voraussetzungen von Lemma 4.2 für ein $\gamma > 0$ erfüllt. Dann existiert ein $\delta > 0$, sodass unter der Bedingung $d_p \left(L_i^{\mathbb{M}}, \hat{L}_i^{\mathbb{M}} \right) < \delta$ der Monte Carlo Fehler für den Eigenmittelkapitalbedarf K unter Verwendung des Value at Risk, $\rho = \text{VaR}_{\alpha}$, beschränkt ist durch

$$\left\| K_i^{\mathbb{M}} - \hat{K}_i^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} \leq \sqrt[4]{\frac{1}{n} M C E_{i, \text{VaR}}^{\mathbb{M}}}$$

mit

$$MCE_{i, VaR}^{\mathbb{M}} = \begin{cases} \sqrt[4]{\frac{1}{n} \sqrt{C_{vv}^{\mathbb{P}}} + \left(1 + \frac{1}{\gamma}\right) \sqrt[4]{\text{tr}(C_{AA}^{\mathbb{P}} E_{\mathbb{P}}[AA^T])}}, & \text{für } i = 1 \text{ und } \mathbb{M} = \mathbb{P}, \\ \left(1 + \frac{1}{\gamma}\right) \sqrt[4]{\text{tr}(C_{AA}^{\mathbb{Q}} E_{\mathbb{P}}[AA^T])}, & \text{für } i = 1 \text{ und } \mathbb{M} = \mathbb{Q}, \\ \left(1 + \frac{1}{\gamma}\right) \sqrt[4]{\text{tr}(C_{BB}^{\mathbb{M}} E_{\mathbb{P}}[BB^T])}, & \text{für } i = 2, \end{cases}$$

wobei $C_{vv}^{\mathbb{M}}$, $C_{AA}^{\mathbb{M}}$ und $C_{BB}^{\mathbb{M}}$ die diagonalen 1×1 -, $|A| \times |A|$ - und $|B| \times |B|$ -Blöcke der Matrix $C^{\mathbb{M}}$ sind.

Beweis. Folgt ähnlich wie im Falle des Expected Shortfalls unter Anwendung der Cauchy-Schwarzschen Ungleichung. Für $i = 1$ gilt

$$\begin{aligned} \left\| K_i^{\mathbb{M}} - \hat{K}_i^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} &= \left\| v^{\mathbb{M}} - \hat{v}^{\mathbb{M}} + \rho\left(\alpha_A^{\mathbb{M}T} A\right) - \rho\left(\hat{\alpha}_A^{\mathbb{M}T} A | \mathcal{G}\right) \right\|_{L^2(\mathbb{M})} \\ &\leq \left\| v^{\mathbb{M}} - \hat{v}^{\mathbb{M}} \right\|_{L^2(\mathbb{M})} + \left\| \rho\left(\alpha_A^{\mathbb{M}T} A\right) - \rho\left(\hat{\alpha}_A^{\mathbb{M}T} A | \mathcal{G}\right) \right\|_{L^2(\mathbb{M})} \\ &\stackrel{\text{Proposition 4.1}}{\leq} \sqrt{\frac{1}{n} \sqrt{C_{vv}^{\mathbb{M}}} + \left(1 + \frac{1}{\gamma}\right) \sqrt{E_{\mathbb{M}}\left[E_{\mathbb{P}}\left[\left(\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}}\right)^T A | \mathcal{G}\right]\right]}} \end{aligned}$$

Cauchy-Schwarzsche UGL,

$$\stackrel{f(x)=x^2 \text{ konvex}}{\leq} \sqrt{\frac{1}{n} \sqrt{C_{vv}^{\mathbb{M}}} + \left(1 + \frac{1}{\gamma}\right) \sqrt[4]{E_{\mathbb{M}}\left[E_{\mathbb{P}}\left[\left(\left(\alpha_A^{\mathbb{M}} - \hat{\alpha}_A^{\mathbb{M}}\right)^T A\right)^2 | \mathcal{G}\right]\right]}}.$$

Damit folgt nun gleich wie im Beweis vorher die Behauptung des Satzes. Der Fall $i = 2$ funktioniert analog. \square

Wie bei der Analyse des Approximationsfehler gesehen wurde, nimmt dieser bei wachsender Anzahl $|\mathcal{P}|$ der Faktoren ab. Der Monte Carlo Fehler hingegen nimmt bei einer steigenden Stichprobenanzahl n ab, und wächst, wie man aus den letzten Sätzen folgern kann, für steigendes $|\mathcal{P}|$ an. Für ein vorgegebenes Berechnungsbudget gibt es eine optimale Aufteilung zwischen n und $|\mathcal{P}|$. Die Stichprobengröße n und $|\mathcal{P}|$ werden so gewählt, dass der totale Schätzfehler minimiert wird. Diese optimale Wahl ist jedoch von Problem zu Problem unterschiedlich.

Die Analyse und Berechnung der replizierenden Portfolio Methode, wurde sowohl für das risikoneutrale als auch für das reale Maß betrachtet. In vielen Punkten hat jedoch die Berechnung über das risikoneutrale Maß Vorteile gegenüber der Berechnung über das reale Maß. Da man mit dem Satz 4.4, eine konsistente Schätzung für die Risikomaße erhält, auch wenn man die Lösung des Minimierungsproblems (63) für den $L^2(\mathbb{Q})$ berechnet, werden die Beispiele im Kapitel 9 alle über das risikoneutrale Maß berechnet.

8.1 Anwendung von Varianzreduzierenden Methoden

Da es Teil des Hauptproblems ist den Erwartungswert $E \left[\begin{pmatrix} AZ \\ BZ \end{pmatrix} \right]$ durch Monte Carlo Simulation zu schätzen, ist es sinnvoll für die Verbesserung der Konvergenz Varianzreduktionsmethoden anzuwenden.

In dieser Arbeit werden in den Beispielen Finanzinstrumente zur Replikation herangezogen, welche von unabhängigen geometrischen Brownschen Bewegungen gelenkt werden. Daher bietet sich die Methoden des Stratified Samples bzw. bei Pfadabhängigkeit Latin Hypercube Sample und die kontrollierende Zufallsvariablen an. Die Anwendung dieser Methoden kann wie folgt vorgenommen werden.

- Kontrollierende Zufallsvariable:

Da für $\mathbb{M} = \mathbb{Q}$ gilt, dass $E_{\mathbb{Q}}[A] = 0$ und $E_{\mathbb{Q}}[B] = 0$ sowie vorausgesetzt wird, dass $E_{\mathbb{Q}}[Z] = 0$ ist, ist es möglich diese Zufallsvariablen als Kontrolle zu verwenden. Es wird hier gezeigt, wie man Z als Kontrolle anwenden kann. Das wird auch später in den Beispielen verwendet.

Seien

$$\text{diag}(\hat{a}_i^n) Z$$

und

$$\text{diag}(\hat{b}_i^n) Z$$

wobei durch

$$\hat{a}_i^n = \frac{\sum_{j=1}^n (A_i^{(j)} Z^{(j)} - \overline{A_i Z}) (Z^{(j)} - \bar{Z})}{\sum_{j=1}^n (Z^{(j)} - \bar{Z})^2}$$

der Schätzer für den optimale Koeffizient dieser Methode berechnet wird. \hat{b}_i^n kann analog berechnet werden, dann kann man statt (75)

$$\hat{v}^{\mathbb{Q}} = 0, \begin{pmatrix} \hat{\alpha}_A^{\mathbb{Q}} \\ \hat{\alpha}_B^{\mathbb{Q}} \end{pmatrix} = \mathcal{N}^{-1} \frac{1}{n} \sum_{j=1}^n \left(\begin{pmatrix} A^{(j)} Z^{(j)} \\ B^{(j)} Z^{(j)} \end{pmatrix} - \begin{pmatrix} \text{diag}(\hat{a}_i^n) Z^{(j)} \\ \text{diag}(\hat{b}_i^n) Z^{(j)} \end{pmatrix} \right) \quad (80)$$

berechnen.

- Stratified Sampling:

Es wird davon ausgegangen, dass die Finanzderivate welche repliziert werden sollen von einer Brownschen Bewegung bezüglich $\mathbb{M} = \mathbb{Q}$ gelenkt werden. Daher kann man für die Reduktion der Varianz die Methode Stratified Sampling bzw. Latin Hypercube Sampling verwenden. Will man ein Finanzderivat X replizieren, welches nur vom Endwert

abhängt kann man die Methode (5.4.4) verwenden um von den Pfaden dieser simulierten Stichprobe von Brownschen Bewegungen den gesuchten Erwartungswert zu schätzen.

Sucht man das Risikomaß eines Finanzderivates X , welches pfadabhängig ist, wie zum Beispiel bei exotischen Optionen, dann kann man ein Latin Hypercube Sample wie in (43) als Grundlage für die Replikation verwenden.

Im nächsten Kapitell werden Beispiele betrachtet, anhand jener unter anderem diese Methoden miteinander verglichen werden.

9 Beispiele

Es werden nun ein ein paar Beispiele besprochen, bei denen man die Anwendung des dynamischen replizierenden Portfolios mit dem des statischen replizierenden Portfolios vergleicht. Dabei wird auch der Einfluss der Varianzreduktionsmethoden betrachtet.

Um einen wirklichen Vergleich der Konvergenz zu erhalten, werden ausschließlich Portfolios betrachtet, welche Finanzderivate enthalten, für deren Einjahresgewinn eine explizite Bepreisungsformel existiert.

Im ersten Beispiel wird eine geometrische Brownsche Bewegung durch eine Brownsche Bewegung mit Drift, welche unter \mathbb{Q} eine Brownsche Bewegung ist repliziert. Für eine Brownsche Bewegung kann man die Gramsche Matrix einfach berechnen, da sie sich durch die Unabhängigkeit der Zuwächse auf eine Diagonalmatrix reduziert.

In den folgenden Beispielen werden zuerst einzelne Finanzderivate betrachtet, welche auch nur von einem stochastischen Prozess gelenkt werden. Dieser ist durch eine geometrische Brownsche Bewegung gegeben. Daher werden in diesen ersten Beispielen zur Replikation jeweils nur ein Finanzinstrument, welches durch eine Brownsche Bewegung gelenkt wird, verwendet. Es werden die Methoden des statisch replizierenden Portfolios mit dem des dynamischen replizierenden Portfolios, sowie die Anwendung der beiden Varianzreduktionsmethoden, welche im vorigen Kapitel beschrieben wurden, verglichen.

Das letzte Beispiel ist ein Portfolio, welches aus zwei Finanzderivaten besteht, welche jeweils von unabhängigen geometrischen Brownschen Bewegungen gelenkt werden. In diesem Beispiel werden zwei dynamischen replizierende Portfolios, für welche unterschiedliche Basisfunktionen verwendet werden, unter Anwendung der Varianzreduktionsmethode Stratified Sample bzw. Latin Hypercube Sample verglichen. Die beiden Basen unterscheiden sich wie folgt voneinander:

In beiden Fällen werden die Basisfunktionen, wie in 7.1 beschrieben erstellt. Eine Basis besteht aus allen möglichen Termen beinhaltet also auch die gemischten Terme der Finanzinstrumente. Die andere wird in dieser Hinsicht eingeschränkt, dass sie keine gemischten Terme enthält.

Die Annahmen, welche jetzt getroffen werden, gelten für alle folgenden Beispiele. Die Laufzeit der Finanzderivate wird mit $T = 5$ festgelegt, und der Modellhorizont wird in $0 = t_0 < t_1 < \dots < t_n = T$ äquidistante Punkte eingeteilt, wobei $t_i - t_{i-1} = \frac{1}{4}$, was bedeutet, dass es möglich ist das Portfolio quartalsweise auszubalancieren.

Die Finanzinstrumente werden von unabhängigen Brownschen Bewegungen mit Drift γ_i für $i = 1, \dots, m$ gelenkt, das heißt die Finanzinstrumente werden von $G_t = \gamma t + W_t$ gelenkt, wobei $(W_t)_{0 \leq t \leq T}$ eine m -dimensionale Brownsche Bewegung bezüglich \mathbb{P} und $\gamma \in \mathbb{R}^m$ ein konstanter Vektor ist.

γ beschreibt den Unterschied zwischen den zwei Maßen und wird als market price of risk bezeichnet. Der Maßwechsel zum risikoneutralen Maß \mathbb{Q} , wird nun über die Dichte

$$D_t = \exp\left(-\frac{\|\gamma\|^2}{2}t - \gamma^T W_t\right)$$

definiert. Unter dem risikoneutralen Maß ist G_t nach dem Satz von Girsanov eine Brownsche Bewegung.

Die folgenden Beispiele werden alle unter dem risikoneutralen Maß betrachtet, da sich unter dem Maß \mathbb{Q} die Berechnung der Koeffizienten α vereinfacht. Für die Berechnung der Risikomaße wird der Satz 4.4 herangezogen. In den ersten drei Beispielen wurden jeweils 50 Durchläufe mit je $n = 5000$ Stichproben simuliert. Im letzten Beispiel wurden 10 Durchläufe zu je $n = 5000$ Stichproben gestartet. Der Durchschnitt dieser Durchläufe wird für die Analyse herangezogen. γ wird in allen Simulationen mit 0.2 festgelegt. Es wird in den Beispielen der Schätzer L_1 wie er in (68) definiert wurde, betrachtet. $|\mathcal{J}|$ wird von $1, \dots, 4$ festgelegt, da für eine größere Anzahl von \mathcal{J} sich für die Basis, welche zur Berechnung des Schätzers L herangezogen wird, nichts mehr ändert, was daher kommt, dass das Ausbalancieren quartalsweise erlaubt ist, und daher für A nur Polynome bis zum Grad vier verwendet werden. Alle Polynome höheren Grades, fallen in den Vektor B . Im Beispiel des Portfolios bestehend aus zwei Finanzderivaten betrachtet man $|\mathcal{J}|$ bis 3, da der rechnerische Aufwand eines dynamischen pfadabhängigen Portfolios, welcher Basisfunktionen höheren Grades verwendet, die Kapazitäten von Matlab überschreitet. Natürlich wäre es auch möglich, die Zeitpunkte, an denen es erlaubt ist das Portfolio auszubalancieren zu reduzieren, was jedoch bei pfadabhängigen Finanzderivaten, nicht sinnvoll wäre. Des weiteren kann aus den ersten Beispielen gefolgert werden, dass ab $|\mathcal{J}| = 3$ keine signifikante Verbesserung des Fehlers mehr beobachtet werden kann. Dies wird aber später noch genauer analysiert.

Im ersten Beispiel werden die exakten Werte für die beiden Risikomaße berechnet, jedoch werden in den anderen Beispielen Werte betrachtet, welche durch die simulierten Risikomaße, die durch die jeweils angeführte Bepreisungsformeln entstehen, normiert werden. Dies hat den Zweck, dass der prozentuale Fehler besser analysiert werden kann.

9.1 Geometrische Brownsche Bewegung

Als erstes Beispiel wird eine geometrische Brownsche Bewegung repliziert. Es sei durch

$$X = \exp\left(\sigma G_T - \frac{\sigma^2}{2}T\right) - 1$$

die Auszahlungsfunktion zum Zeitpunkt T gegeben, wobei der Wert zum Zeitpunkt t_0 gleich 1 ist. In diesem Fall gilt $m = 1$ und der Preisprozess des

Finanzinstrumentes G agiert wie eine Brownsche Bewegung mit Drift, dass heißt

$$G_t = \gamma t + W_t,$$

wobei $W = (W_t)_{0 \leq t \leq T}$ eine Brownsche Bewegung bzgl. \mathbb{P} ist, und γ ist der market price of risk, welcher vorher bereits festgelegt wurde. Der Dichtprozess für den Maßwechsel ergibt sich in diesem Fall wie folgt:

$$D_t = \exp\left(-\frac{\gamma^2}{2}t - \gamma W_t\right),$$

wobei sich der Preisprozess G unter \mathbb{Q} wie eine Brownsche Bewegung verhält. Da die geometrische Brownsche Bewegung bezüglich dem risikoneutralen Maß ein Martingal ist, erhält man für die Wertänderung nach einem Jahr $L = \exp\left(\sigma G_1 - \frac{\sigma^2}{2}\right) - 1$.

Es wird also versucht, die Geometrische Brownsche Bewegung als Linearkombination von multilinearen Polynomen bis zum Grad vier, der Zuwächse der Brownschen Bewegung darzustellen. In der Simulation wird $\sigma = 0.2$ gesetzt.

Da für die geometrische Brownsche Bewegung die Risikomaße leicht berechnet werden können, wird die Schätzung der Risikomaße mit den exakten Werten verglichen. Nachdem der Einjahergewinn von X eine geometrische Brownsche Bewegung zum Zeitpunkt $t = 1$ abzüglich einer Konstanten ist, kann man den Value at Risk laut Definition und in Anwendung der Eigenschaften der Quantilfunktionen, welche im Satz 2.4 und dem Korollar 2.1 bewiesen wurden ohne größeren Aufwand für dieses Beispiel berechnen.

Für den Value at Risk gilt

$$\begin{aligned} VaR_\alpha(L) &= -q_L^+(1 - \alpha) \\ &= -\exp\left(\sigma\gamma - \frac{\sigma^2}{2} + \sigma q_X^+(1 - \alpha)\right) + 1 \stackrel{\alpha=0.995}{=} 0.3905, \end{aligned}$$

wobei X eine standardnormalverteilte Zufallsvariable ist. Die Berechnung des Expected Shortfalls ist etwas aufwendiger, da von einer stetigen Verteilung ausgegangen wird, gilt

$$ES_\alpha(L) = -E[L|L \leq -VaR_\alpha(L)]$$

Verwendet man die Darstellung des Value at Risk von eben vorhin und die Tatsache, dass $\mathbb{P}(L \leq -VaR_\alpha(L)) = 1 - \alpha$, kann man den Expected Shortfall durch

$$ES_\alpha(L) = -\frac{1}{1 - \alpha} \left(1 - \Phi(\lambda - q_X^+(1 - \alpha))\right) + 1,$$

ausgedrückt werden, wobei hier X wieder eine standardnormalverteilte Zufallsvariable ist. Für $\alpha = 0.99$ erhält man

$$ES_\alpha(L) = 0.42377.$$

In diesem Beispiel werden statisches replizierendes Portfolio, dynamisches replizierendes Portfolio, kontrollierende Zufallsvariable angewendet auf das dynamische replizierende Portfolio und Terminal Stratification angewendet auf das dynamische replizierende Portfolio miteinander verglichen.

Wobei als kontrollierende Zufallsvariable Z herangezogen wird. Beim Terminal Stratification Sample wird die reelle Achse in 5000 Teilmengen eingeteilt, welche alle Wahrscheinlichkeit 0.0002 haben, und aus jeder Schicht wird eine Stichprobe simuliert.

In der ersten Tabelle 9.1 sind die Werte der quadrierten Korrelationskoeffizienten von dem replizierenden Portfolio V und dem Portfolio Z und in der zweiten 9.1 die von dem simulierten Einjahresverlust L und dem von L_1 angegeben.

R^2	$ \mathcal{J} = 1$	$ \mathcal{J} = 2$	$ \mathcal{J} = 3$	$ \mathcal{J} = 4$
statisches RP	0.9035	0.9035	0.9035	0.9035
dynamisches RP	0.9004	0.9464	0.8018	0.5244
mit kontrollierender ZV	0.9007	0.9468	0.8024	0.5270
mit Terminal Stratification	0.9008	0.9472	0.8022	0.5188

Tabelle 3: Korrelationskoeffizienten von V und Z

R^2	$ \mathcal{J} = 1$	$ \mathcal{J} = 2$	$ \mathcal{J} = 3$	$ \mathcal{J} = 4$
statisches RP	0.9800	0.9800	0.9800	0.9800
dynamisches RP	0.9772	0.9836	0.9782	0.9762
mit kontrollierender ZV	0.9772	0.9836	0.9789	0.9763
mit Terminal Stratification	0.9769	0.9834	0.9759	0.9734

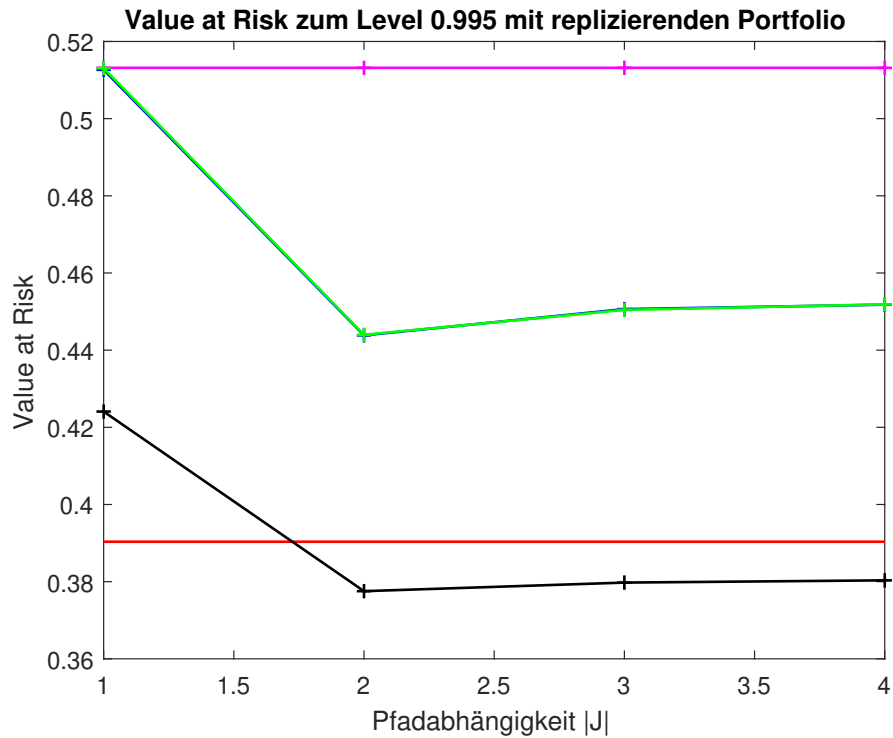
Tabelle 4: Korrelationskoeffizienten von L und L_1

Man kann sehen, dass sich die Korrelationskoeffizienten von Z und V für $|\mathcal{J}| > 2$ verschlechtern, die für L und L_1 aber stabil bleiben. Daran kann man erkennen, dass der Schätzer L_2 wie er in (69) definiert wurde, keine gute Schätzung für $|\mathcal{J}| > 2$ bringen würde. Es sind die Koerrelationskoeffizienten der dynamischen replizierenden Portfolios alle sehr ähnlich. In den folgenden Graphiken sieht man jedoch, dass die Schätzung der Risikomaße beim Stratified Sample im Vergleich mit den anderen besser ist.

In der Graphik 6 werden die verschiedenen Value at Risk Werte für $|\mathcal{J}| = 1, \dots, 4$ angegeben. In der Graphik 7 werden die Werte des Expected Shortfalls angegeben, wobei bei beiden Graphiken die rote Linie der exakte berechnete Wert ist.

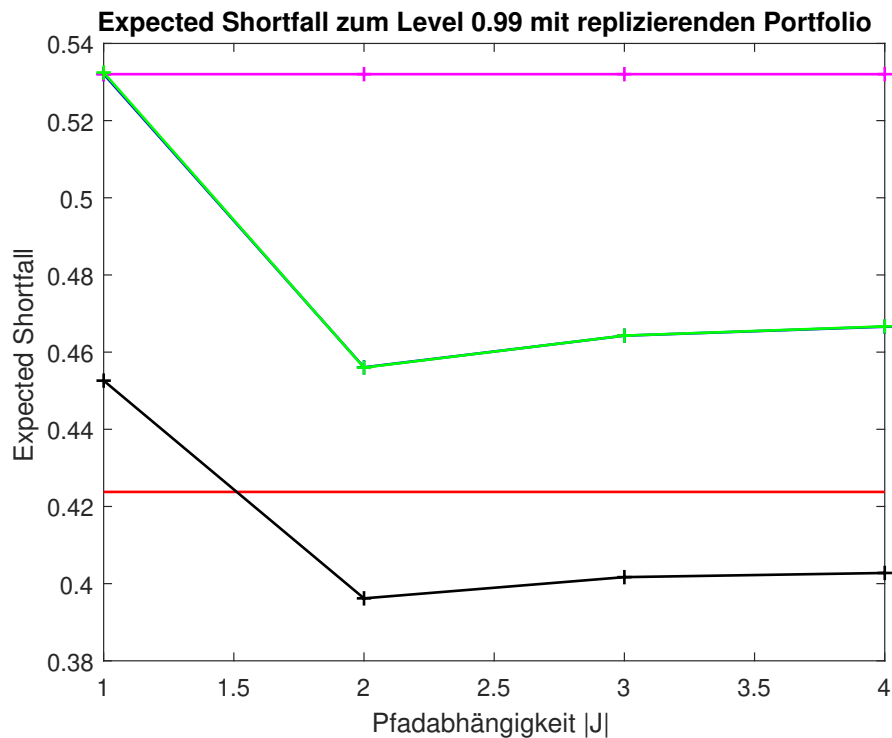
In beiden Fälle kann man sehen, dass das pfadabhängige Portfolio auf alle Fälle bessere Schätzungen als das statische Portfolio bringt. Bei den Vari-

anzreduktionsmethoden sieht man, dass das Stratified Sample für $|\mathcal{J}| > 2$ gute Schätzwerte hervorbringt, wobei sich Schätzwerte für die kontrollierende Zufallsvariable kaum von jenen des replizierenden Portfolios ohne Varianzreduktion unterscheiden. Die Schätzungen des Value at Risk sind etwas genauer als die des Expected Shortfalls.



Die rote Linie ist der exakte Wert des Value at Risk, das statische RP ist in magenta, das dynamische RP ohne Varianzreduktion ist blau, mit kontrollierender Zufallsvariable Z grün, und mit dem Stratified Sample ist es schwarz.

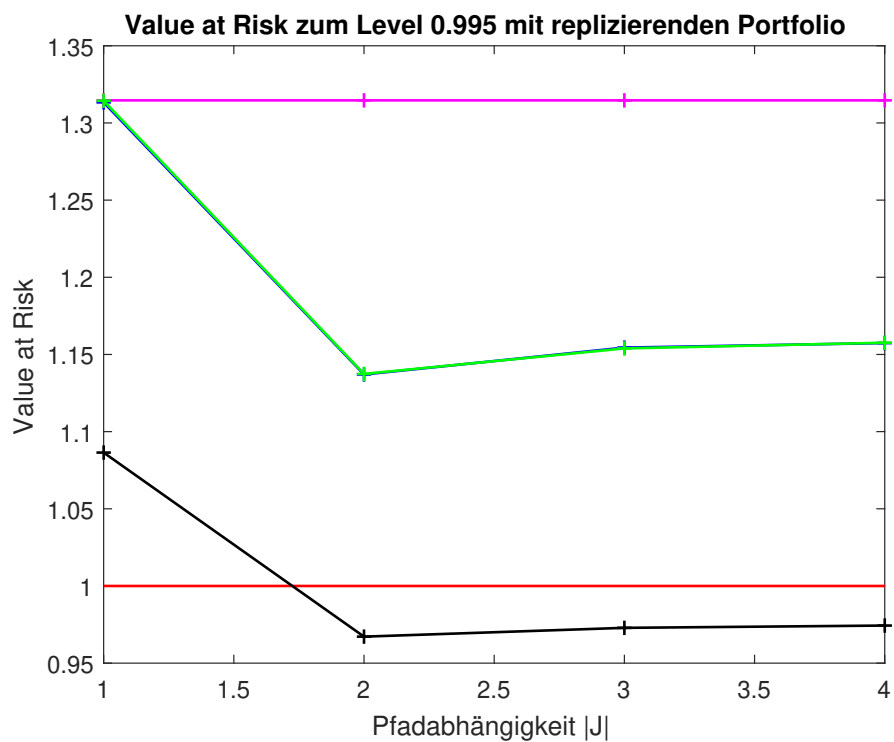
Abbildung 6: Value at Risk Wert mit $\alpha = 0,995$ für $|\mathcal{J}| = 1, \dots, 4$



Die rote Linie ist der exakte Wert des Expected Shortfalls, das statische RP ist in magenta, das dynamische RP ohne Varianzreduktion ist blau, mit kontrollierender Zufallsvariable Z grün, und mit dem Stratified Sample ist es schwarz.

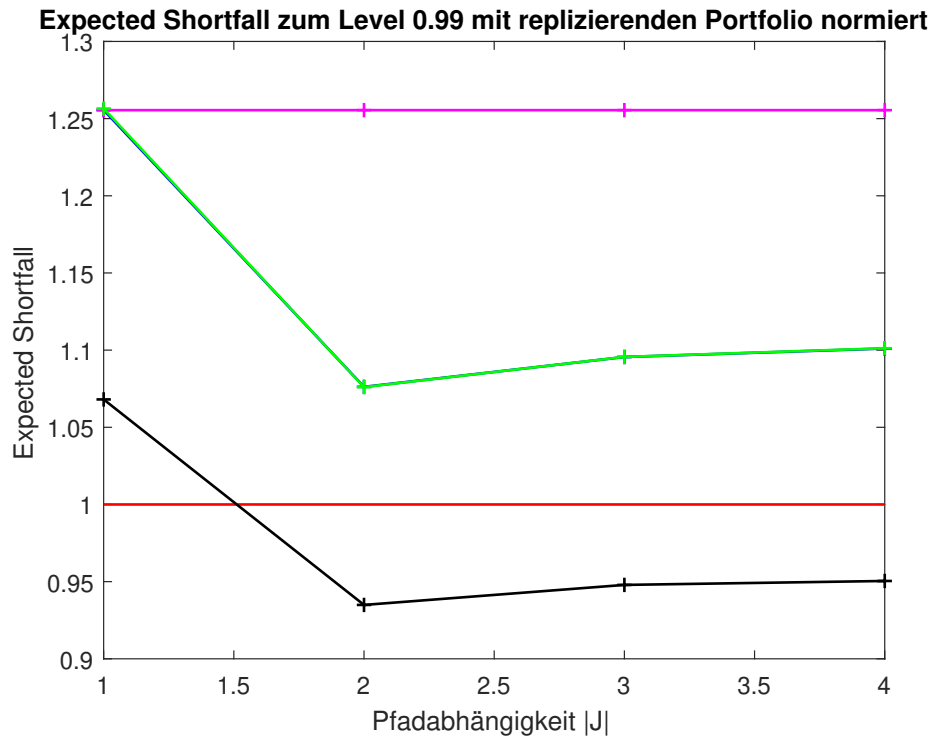
Abbildung 7: Expected Shortfall mit $\alpha = 0.99$ für $|\mathcal{J}| = 1, \dots, 4$

In den folgenden beiden Graphiken 8 und 9, betrachtet man die normierten Werte der beiden Risikomaße. Man kann für das Replizierende Portfolio mit Terminal Stratification gut erkennen, dass sich der Fehler beim Value at Risk von über 10 % auf unter 5 % verbessert. Der Fehler der anderen Portfolios bleibt leicht über 10 %. Im Gegensatz dazu verbessert sich der Fehler des Stratified Samples beim Expected Shortfall auf nur ca. 5 %, wobei hier die anderen sich auf etwas unter 10 % verbessern. Die Verbesserung des Fehlers ist für $|\mathcal{J}| = 4$ nicht mehr signifikant. Da die Anzahl der Basisfunktionen so klein wie möglich gehalten werden soll, wäre hier $|\mathcal{J}| = 3$ die perfekte Wahl.



Die rote Linie ist der normiert Wert des Value at Risk, das statische RP ist in magenta, das dynamische RP ohne Varianzreduktion ist blau, mit kontrollierender Zufallsvariable Z grün, und mit dem Stratified Sample ist es schwarz.

Abbildung 8: Value at Risk Wert mit $\alpha = 0,995$ für $|\mathcal{J}| = 1, \dots, 4$



Die rote Linie ist der normierte Wert des Expected Shortfalls, das statische RP ist in magenta, das dynamische RP ohne Varianzreduktion ist blau, mit kontrollierender Zufallsvariable Z grün, und mit dem Stratified Sample ist es schwarz.

Abbildung 9: Expected Shortfall mit $\alpha = 0.99$ für $|\mathcal{J}| = 1, \dots, 4$

9.2 Europäische Call Option

Im zweiten Beispiel wird eine Europäische Call Option im Black Scholes Modell betrachtet, was bedeutet, dass die Auszahlungsfunktion Z durch

$$Z_T = (S_T - K)^+$$

gegeben ist, wobei $(S_t)_{0 \leq t \leq T}$ eine geometrische Brownsche Bewegung ist, welche durch eine Brownsche Bewegung $(G_t)_{0 \leq t \leq T}$ bezüglich \mathbb{Q} gelenkt wird. Es gilt

$$S_t = S_0 \exp\left(-\frac{\sigma^2}{2}t + \sigma G_t\right).$$

Der Preis der europäischen Call Option zum Zeitpunkt t ist im Black Scholes Modell nach [ABLM13], dann wie folgt gegeben:

$$Z_t = S_t \Phi(d_1) - K \Phi(d_2),$$

wobei

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}$$

und

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Da wir von einer fairen Bepreisung zum Zeitpunkt 0 ausgehen, ist der Einjahresgewinn und der Gewinn zum Endzeitpunkt T durch

$$L = S_1\Phi(d_1) - k\Phi(d_2) - Z_0$$

und

$$X = Z_T - Z_0$$

gegeben.

Für die Simulation dieses Problems seien $S_0 = 100$ und $K = 100$. Die anderen Parameter werden wie im Beispiel zuvor gewählt.

Wie auch schon im vorigen Beispiel wird das Stratified Sample mit je einer Stichproben aus 5000 Schichten berechnet.

In den Tabellen 9.2 und 9.2 sind wieder die quadrierten Korrelationskoeffizienten von Z und V und L und L_1 angegeben.

Die Korrelationskoeffizienten sind im Vergleich zum vorigen Beispiel um einiges niedriger. Die Entwicklung ist im Prinzip gleich. Für $|\mathcal{J}| = 2$ sind sie am größten, danach fallen sie wieder ab. Der Korrelationskoeffizient des statischen Replizierenden Portfolios weist darauf hin, dass die Schätzung für die Risikomaße sehr schlecht sein kann, was die Graphiken auch bestätigen.

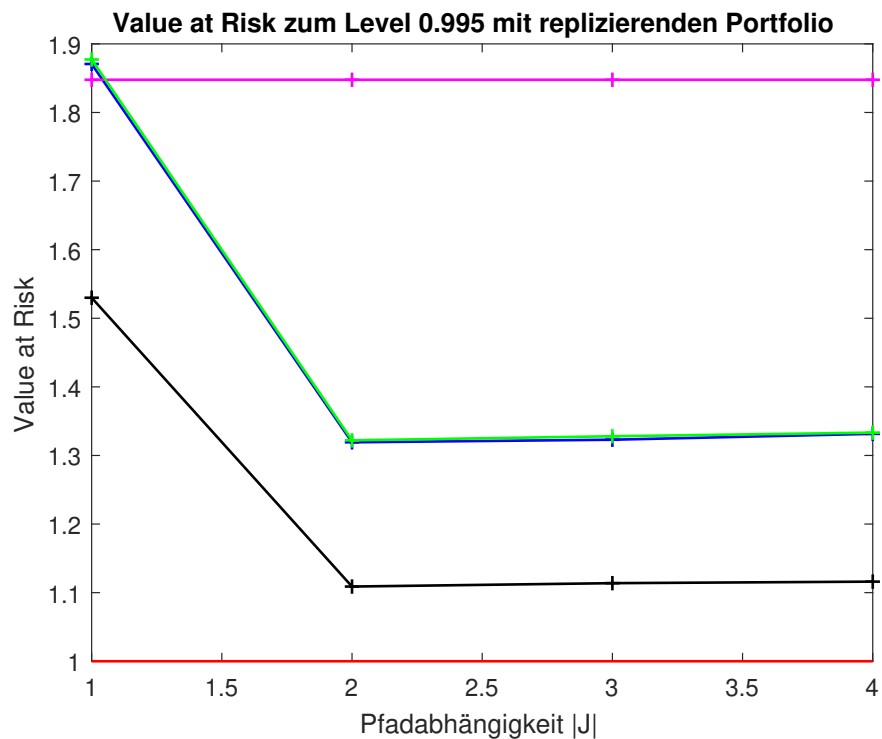
R^2	$ \mathcal{J} = 1$	$ \mathcal{J} = 2$	$ \mathcal{J} = 3$	$ \mathcal{J} = 4$
statisches RP	0.5858	0.5858	0.5858	0.5858
dynamisches RP	0.5853	0.8902	0.7663	0.5064
mit kontrollierender ZV	0.5854	0.8901	0.7648	0.5042
mit Terminal Stratification	0.5866	0.8943	0.7807	0.5199

Tabelle 5: Korrelationskoeffizienten von Z und V

R^2	$ \mathcal{J} = 1$	$ \mathcal{J} = 2$	$ \mathcal{J} = 3$	$ \mathcal{J} = 4$
statisches RP	0.8875	0.8875	0.8875	0.8875
dynamisches RP	0.8830	0.9543	0.9494	0.9472
mit kontrollierender ZV	0.8830	0.9544	0.9495	0.9473
mit Terminal Stratification	0.8831	0.9563	0.9520	0.9495

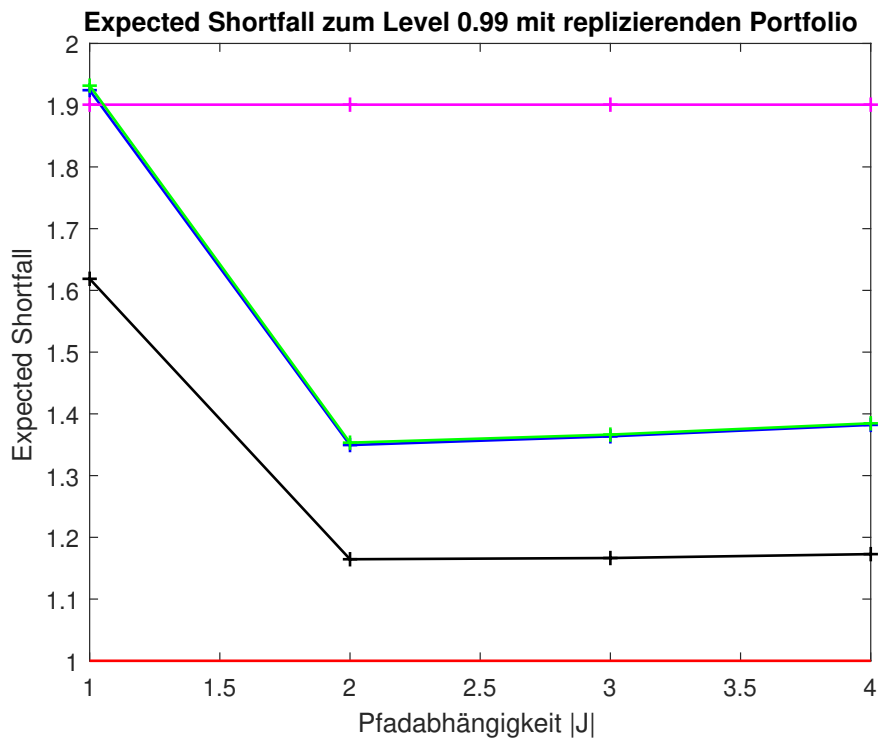
Tabelle 6: Korrelationskoeffizienten von L und L_1

In den folgenden Graphiken sind die normierten Werte der Risikomaße zu sehen. Wie schon aus den Korrelationskoeffizienten ersichtlich ist, ist die Verbesserung des Fehlers von Basisfunktionen, welche nur Polynome des Grades eins verwenden, auf jene, welche zusätzlich auch Polynome des Grades zwei verwenden am gravierendsten. Es gilt wieder, dass das replizierende Portfolio mit dem Stratified Sample den gewünschten Wert am besten approximiert. Jedoch verbessert sich der Schätzfehler in diesem Beispiel nur auf ca. 10 %, worauf auch schon die Korrelationskoeffizienten hinweisen.



Hier ist die rote Linie der normierte Wert des Risikomaßes, welches man durch die Schätzung von L bekommt, wenn man die angegebene Bepreisungsformel verwendet. In Magenta ist wieder das statische RP, in blau das dynamische RP ohne Varianzreduktion, in grün mit kontrollierender Zufallsvariable und in schwarz mit Terminal Stratification angegeben, wobei die Werte durch die Schätzung des Risikomaßes bezüglich L normiert wurden.

Abbildung 10: Value at Risk Wert mit $\alpha = 0,995$ für $|\mathcal{J}| = 1, \dots, 4$



Es ist wieder $K = 1$ in rot, das gewünschte Kapital. In Magenta ist das statische RP, in blau das dynamische RP ohne Varianzreduktion, in grün mit kontrollierender Zufallsvariable und in schwarz mit Terminal Stratification angegeben, wobei die Werte durch die Schätzung des Risikomaßes bezüglich L normiert wurden.

Abbildung 11: Expected Shortfall mit $\alpha = 0.99$ für $|\mathcal{J}| = 1, \dots, 4$

9.3 Asiatische Put Option

In diesem Beispiel wird die Methode des Latin Hypercubes zur Varianzreduktion verwendet. Es wird ein replizierendes Portfolio ohne Varianzreduktion mit einem Portfolio verglichen, welches über ein Latin Hypercube Sample berechnet wird. Da das statische replizierende Portfolio nicht an die Schätzungen eines dynamischen replizierenden Portfolios herankommt, wird in diesem Beispiel darauf verzichtet. Genauso wird auch die Methode der kontrollierenden Zufallsvariable nicht weiter betrachtet, da sie keine wirkliche Verbesserung im Vergleich zur klassischen Monte Carlo Methode bringt. Es wird eine pfadabhängige Asiatische Put Option, welche über das geometrische Mittel definiert wird, repliziert. Die Auszahlungsfunktion einer

solchen Option zum Zeitpunkt T ist durch

$$Z_T = \left(K - \left(\prod_{i=1}^m S_{t_i} \right)^{\frac{1}{m}} \right)^+$$

wobei $(S_t)_{0 \leq t \leq T}$ eine geometrische Brownsche Bewegung mit konstanter Volatilität ist, also $S_t = S_0 \exp\left(-\frac{\lambda^2}{2}t + \lambda G_t\right)$, dabei ist $(G_t)_{0 \leq t \leq T}$ wieder eine Brownsche Bewegung bzgl. \mathbb{Q} und $0 = t_0 < t_1 < \dots < t_m = T$ sind die Tage an denen der Wert der Aktie S für die Bildung des geometrischen Mittels herangezogen wird. Diese sind zum Zeitpunkt $t_0 = 0$ bekannt. Den Preis dieser Option kann man explizit mit Hilfe der Black Scholes Formel für europäische Optionen bestimmen. Diese Bepreisung wird in [Zha09] wie folgt hergeleitet. Es gilt $t_j - t_{j-1} = \Delta t$ für $j = 2, \dots, m$, damit ist $T - t_k = (m - k) \Delta t$ für $k = 1, \dots, m - 1$, also $\Delta t = \frac{T - t_k}{m - k}$. Weiter gilt

$$\left(\prod_{i=1}^m S_{t_i} \right)^{\frac{1}{m}} = \left(\frac{S_{t_m}}{S_{t_{m-1}}} \left(\frac{S_{t_{m-1}}}{S_{t_{m-2}}} \right)^2 \dots \left(\frac{S_{t_{k+1}}}{S_{t_k}} \right)^{m-k} \underbrace{S_{t_k}^{m-k} \prod_{i=1}^k S_{t_i}}_{=: A} \right)^{\frac{1}{m}} = (\times),$$

und da man

$$\frac{S_{t_j}}{S_{t_{j-1}}} = \exp\left(-\frac{\lambda^2}{2}\Delta t + \lambda\sqrt{\Delta t}Z_j\right)$$

mit $Z_j \sim N(0, 1)$ schreiben kann für $j = 2, \dots, m$, erhält man durch einsetzen für die Δt und herausheben der konstanten Faktoren

$$(\times) = \exp\left(\frac{1}{m} \left(\lambda\sqrt{\frac{T - t_k}{m - k}} \sum_{i=1}^{m-k} iZ_i - \frac{\lambda^2}{2} \frac{T - t_k}{m - k} \sum_{i=1}^{m-k} i \right)\right) A^{\frac{1}{m}} = (\times \times).$$

Es gilt, dass

$$Y = \lambda\sqrt{\frac{T - t_k}{m - k}} \sum_{i=1}^{m-k} iZ_i \sim N\left(0, \lambda^2 \frac{T - t_k}{m - k} \sum_{i=1}^{m-k} i^2\right),$$

und die beiden Summen können durch

$$\sum_{i=1}^{m-k} i = \frac{(m - k)(m - k + 1)}{2}$$

und

$$\sum_{i=1}^{m-k} i^2 = \frac{(m - k)(m - k + 1)(2(m - k) + 1)}{6}.$$

berechnet werden. Damit erhält für den obigen Ausdruck

$$(\times \times) = A^{\frac{1}{m}} \exp \left(\left(\mu - \frac{\bar{\lambda}^2}{2} \right) (T - t_k) + \bar{\lambda} \sqrt{(T - t_k)} X \right),$$

wobei X eine standardnormalverteilte Zufallsvariable ist, und für die Parameter gilt

$$\bar{\lambda}^2 = \frac{\lambda^2 (m - k + 1) (2(m - k) + 1)}{2 \cdot 6}$$

und

$$\mu = \frac{\bar{\lambda}^2}{2} - \frac{\lambda^2 (m - k + 1)}{2 \cdot 2m}.$$

Dabei ist μ so definiert, dass

$$\mu - \frac{\bar{\lambda}^2}{2} = -\frac{\lambda^2 (m - k + 1)}{2 \cdot 2m}.$$

Damit kann man den Preis der Asiatischen Put-Option mit der Black Scholes Formel für europäische Optionen durch

$$P(\tilde{S}_{t_k}, K) = K \Phi(-d_-) - \tilde{S}_{t_k} \exp(\mu(T - t_k)) \Phi(-d_+)$$

wobei

$$d_1 = \frac{\ln\left(\frac{\tilde{S}_{t_k}}{K}\right) + \left(\mu + \frac{\bar{\lambda}^2}{2}\right)(T - t_k)}{\bar{\lambda} \sqrt{T - t_k}}$$

und

$$d_2 = d_1 - \bar{\lambda}(T - t_k)$$

mit

$$\tilde{S}_{t_k} = A^{\frac{1}{m}}$$

berechnen. Die die Auszahlung am Ende der Laufzeit ist damit durch

$$X = \left(K - \left(\prod_{i=1}^m S_{t_i} \right)^{\frac{1}{m}} \right)^+ - Z_0$$

festgelegt, und der Einjahresgewinn kann durch

$$L = K \Phi(-d_2) - S_1 \exp(\mu(T - 1)) \Phi(-d_1) - Z_0$$

berechnet werden.

In diesem Beispiel wird eine Short Position betrachtet, was bedeutet, dass die betrachtete Gewinnfunktion genau das Negative von jener ist, die gerade beschrieben wurde.

Die Parameter für die Simulation sind $K = 80$, $S_0 = 80$ und $\lambda = 0.1$. Für die Partition des Zeithorizontes wählt man die Teilstücke so, dass die mit den Zeitpunkten der Option übereinstimmen, welche man für die Berechnung des geometrischen Mittels heranzieht.

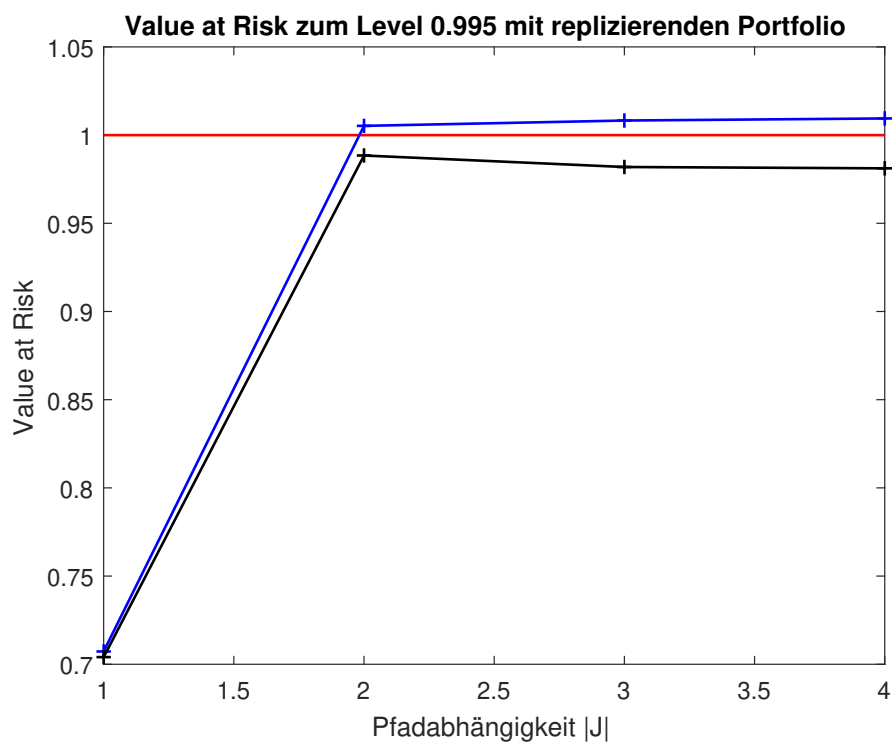
Für das Latin Hypercube Sample werden ein Latin Hypercube Sample mit 5000 Stichproben erstellt.

In der Tabelle 9.3 werden die quadrierten Korrelationskoeffizienten von Z und V angegeben. Es werden hier nur diese Korrelationskoeffizienten betrachtet, da für Finanzderivate, für welche keine Bepreisungsformel angegeben werden kann, ist es auch nur möglich diese zu berechnen. Es werden in diesem Beispiel das dynamische replizierende Portfolio mit einem dynamischen replizierenden Portfolio mit Latin Hypercube Sample verglichen. Die Werte der Korrelationskoeffizienten sind weder für das eine noch für das andere Portfolio besser. Betrachtet man die Graphiken, wird man sehen, dass das Latin Hypercube Sample nicht besser abschneidet, als das replizierende Portfolio ohne Varianzreduktion.

R^2	$ \mathcal{J} = 1$	$ \mathcal{J} = 2$	$ \mathcal{J} = 3$	$ \mathcal{J} = 4$
dynamisches RP	0.7744	0.9049	0.7666	0.5019
mit Latin Hypercube Sample	0.7744	0.9073	0.7658	0.5110

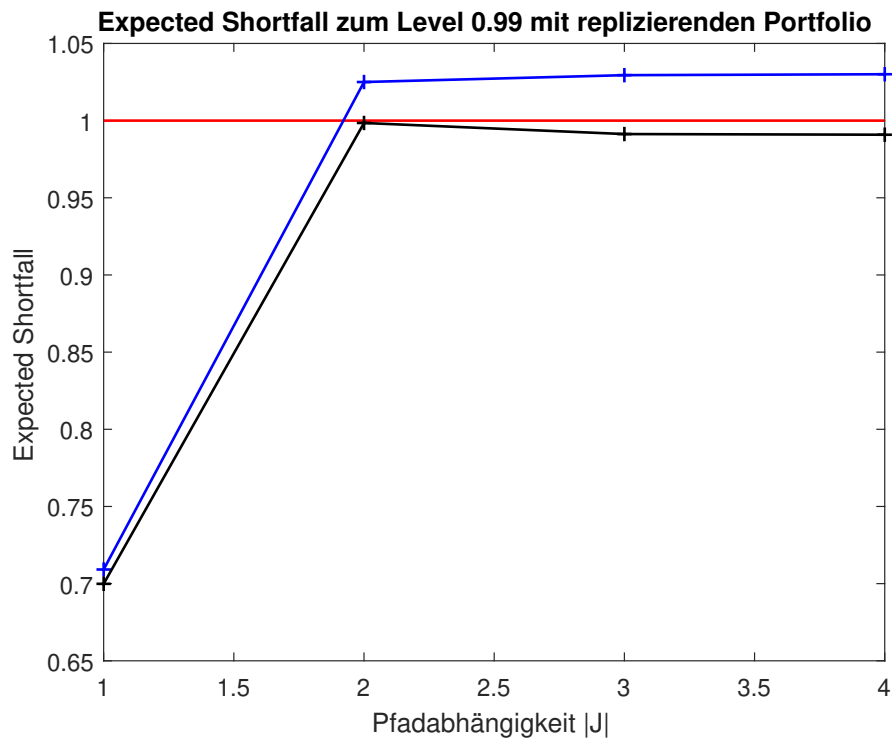
Tabelle 7: Korrelationskoeffizienten von Z und V

In den Graphiken der Risikomaße 12 und 13 sieht man, dass beide replizierenden Portfolios für $|\mathcal{J}| \geq 2$ sehr gute Schätzungen liefern, obwohl die Korrelationskoeffizienten nur leicht besser sind, als die im Beispiel zuvor. Des weiteren kann man beobachten, dass das Latin Hypercube Sample keine Verbesserung bringt. Die beiden replizierenden Portfolios können als gleich bewertet werden.



In blau sind die Werte des replizierenden Portfolios ohne Varianzreduktion abgebildet und in schwarz jene mit Latin Hypercube Sample.

Abbildung 12: Value at Risk Wert mit $\alpha = 0.995$ für $|\mathcal{J}| = 1, \dots, 4$



In blau sind die Werte des replizierenden Portfolios ohne Varianzreduktion abgebildet und in schwarz jene mit Latin Hypercube Sample.

Abbildung 13: Expected Shortfall mit $\alpha = 0.99$ für $|\mathcal{J}| = 1, \dots, 4$

9.4 Portfolio bestehend aus zwei Finanzderivaten

Im letzten Beispiel werden die beiden vorigen Beispiele kombiniert. Man betrachtet ein Portfolio bestehend aus einer Europäischen Call Optionen (Long Position) und einer Asiatischen Put Option (Short Position). Es wird davon ausgegangen, dass die Brownschen Bewegungen, welche die beiden Optionen lenken, also die für die Basisfunktionen herangezogen werden, unabhängig sind. Da in den letzten Beispielen die Einsicht gewonnen wurde, dass die Methode des Stratified Samples sehr effektiv war, und die des Latin Hypercube Samples zumindest genauso effektiv, wie ohne Varianzreduktion, wird in diesem Beispiel nur das replizierende Portfolio mit Anwendung von Latin Hypercube und Terminal Stratification betrachtet, was bedeutet, dass die Brownsche Bewegung, welche die Asiatische Put Option lenkt, über ein Latin Hypercube Sample generiert wird, und die andere wird über Terminal Stratification simuliert.

Hier werden für das Stratified Sample 5000 Schichten zu je einer Stichprobe verwendet, und es wird auch ein Latin Hypercube mit 5000 Stichproben simuliert.

Verglichen werden zwei Replikationen, welche auf unterschiedliche Basisfunktionen aufbauen. Bei der einen Basis werden alle möglichen Basisfunktionen miteinbezogen, wobei bei der zweiten Basis jene Basisterme weglassen werden, welche gemischte Terme der beiden Finanzinstrumente beinhalten, was bedeutet, dass die Basisfunktionen der letzten beiden Beispiele einfach zu einem Satz von Basisfunktionen zusammengefasst wird.

Die Parameter werden direkt aus den Beispielen davor übernommen. Für den Maßwechsel gilt nun,

$$D_t = \exp(-\gamma^2 t - \gamma(W_t^1 + W_t^2)),$$

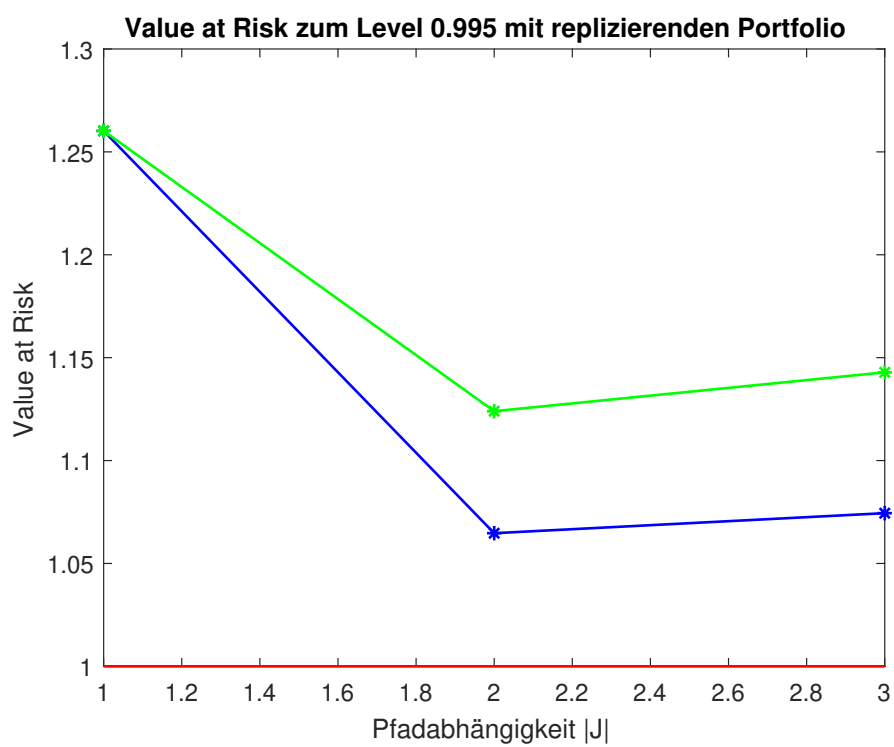
wobei W^1 und W^2 zwei unabhängige Brownsche Bewegungen bezüglich \mathbb{P} sind, sodass für $t \in [0, T]$ $G_t^1 = W_t^1 + \gamma t$ die Brownsche Bewegung ist, welche die Europäische Option lenkt, und $G_t^2 = W_t^2 + \gamma t$ ist jene Brownsche Bewegung, welche die Asiatische Option lenkt. Sowohl G^1 als auch G^2 sind Brownsche Bewegungen bzgl. \mathbb{Q} . Es werden hier nur die Fälle $|\mathcal{J}| \leq 3$ betrachtet, da die Anzahl der Basisfunktionen, vor allem für jenes Portfolio, für welches alle möglichen Basisterme zur Anwendung kommen, schnell steigt, was die rechnerische Effektivität negativ beeinflusst, außerdem hat man in den anderen Beispielen gesehen, dass von $|\mathcal{J}| = 3$ auf $|\mathcal{J}| = 4$ keine signifikante Verbesserung der Schätzung erreicht wurde.

Im folgenden wird mit RP_1 jenes replizierende Portfolio mit der eingeschränkten Anzahl an Basisfunktionen bezeichnet und mit RP_2 das andere. In der Tabelle 9.4 kann man sehen, dass die Korrelationskoeffizienten von RP_1 für $|\mathcal{J}| \geq 2$ besser sind als die Korrelationskoeffizienten von RP_2 .

R^2	$ \mathcal{J} = 1$	$ \mathcal{J} = 2$	$ \mathcal{J} = 3$
RP_1	0.5850	0.8664	0.7477
RP_2	0.5919	0.8354	0.7090

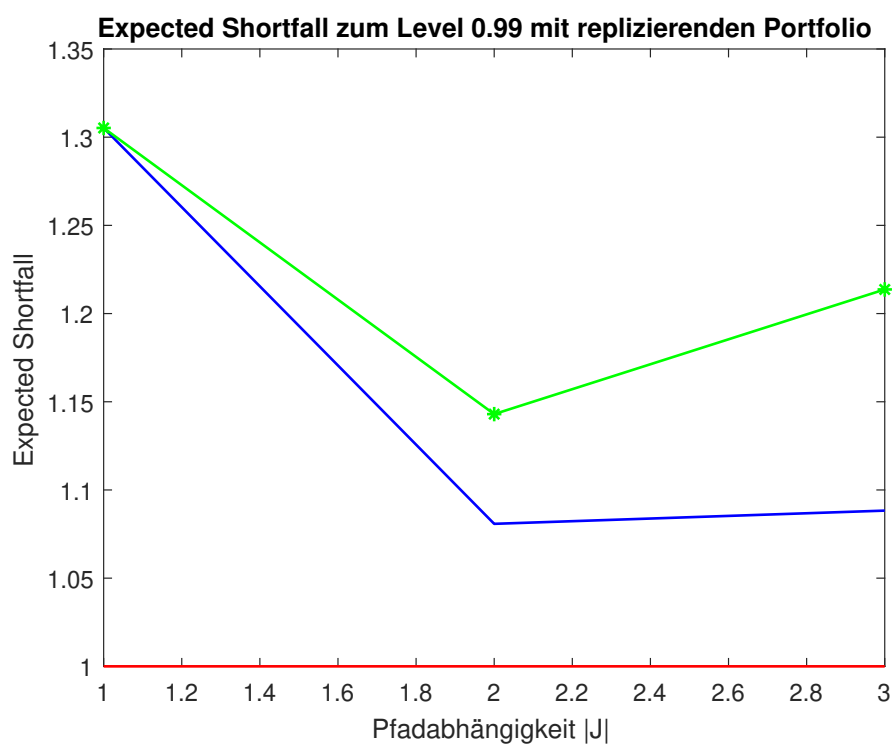
Tabelle 8: Korrelationskoeffizienten von Z und V

Die Ergebnisse der Korrelationskoeffizienten werden durch die Graphiken unterstützt. In den beiden Graphiken 14 und 15 ist ersichtlich, dass die Schätzer jenes Portfolios, welches als Basis die gemischten Terme nicht enthält, besser sind als die anderen Schätzer. Auch hier ist ersichtlich, dass es keine Verbesserung des Fehler für $|\mathcal{J}| > 2$ gibt.



In Grün ist RP_2 und in blau RP_1 abgebildet.

Abbildung 14: Value at Risk Wert mit $\alpha = 0,995$ für $|\mathcal{J}| = 1, \dots, 4$



In Grün ist RP_2 und in blau RP_1 abgebildet.

Abbildung 15: Expected Shortfall mit $\alpha = 0.99$ für $|\mathcal{J}| = 1, \dots, 4$

10 Conclusio

In dieser Arbeit wurde die Schätzung der Risikomaße Value at Risk und Expected Shortfall eines Portfolio behandelt. Dazu hat man mehrere Methoden des replizierenden Portfolios betrachtet, welche eine statische und dynamische Replikation waren. Vom Rechenaufwand aus betrachtet, ist das statische replizierende Portfolio dem anderen überlegen. Jedoch ist der Approximationsfehler dieses Portfolios sehr groß, sodass man keine gute Schätzung für die Risikomaße erhält.

Im Kapitel 7.1 wurde die Konstruktion einer Basis vorgestellt, mit welcher ein dynamisches replizierendes Portfolio berechnet werden kann. Jedoch wird der Rechenaufwand bei der dynamischen Replikation schnell sehr enorm, vor allem wenn ein Portfolio viele Finanzderivate beinhaltet. In den meisten Beispielen wurden daher Finanzderivate betrachtet, für welche nur ein Finanzinstrument für die Berechnung der Basis verwendet wurde.

Im letzten Beispiel wurde ein Portfolio betrachtet, welches durch die Verwendung von zwei Finanzinstrumenten repliziert werden sollte. Dort hat man gesehen, dass es nicht sinnvoll ist eine Basis zu verwenden, welche alle möglichen Terme, der Methode in 7.1 verwendet. Die Methode in welcher man die Basisfunktionen so einschränkt, dass keine Mischterme verwendet werden, was bedeutet, dass für jedes zugrundeliegende Finanzinstrument unabhängig von den anderen eine Basis erstellt wird und diese dann zu einer zusammengefasst werden. Im letzten Beispiel hat diese Methode bessere Schätzer hervorgebracht, als jene wo willkürlich alle Möglichkeiten verwendet wurden. Auch der rechnerische Aufwand ist geringer. Dieses replizierende Portfolio schätzt die Risikomaße auch viel besser als eine statische Replikation, was man aus den Beispielen folgern kann, bei welchen man nur ein Finanzinstrument für die Replikation verwendet hat.

Des weiteren wurde in den Beispielen gesehen, dass es in keinem ab $|\mathcal{J}| = 3$ eine signifikante Verbesserung des Fehlers gegeben hat. Daher wären Basisfunktionen mit Polynomen vom Grad zwei oder drei zu empfehlen, da der Rechenaufwand für die Schätzung bei diesem Grad der Polynome auch noch vertretbar ist.

Abschließend ist noch zu erwähnen, dass von den Varianzreduktionsmethoden, nur die Methode des Stratified Samples eine wirklich erkennbare Verbesserung brachte. Daher ist die Anwendung dieser Methode auf alle Fälle zu empfehlen.

Abbildungsverzeichnis

1	Standardnormalverteilte Stichproben	58
2	Terminal Stratification Brownsche Bewegung	59
3	Nested Simulation	66
4	Least Square Monte Carlo	68
5	Replizierendes Portfolio	71
6	Value at Risk Wert mit $\alpha = 0,995$ für $ \mathcal{J} = 1, \dots, 4$	95
7	Expected Shortfall mit $\alpha = 0,99$ für $ \mathcal{J} = 1, \dots, 4$	96
8	Value at Risk Wert mit $\alpha = 0,995$ für $ \mathcal{J} = 1, \dots, 4$	97
9	Expected Shortfall mit $\alpha = 0,99$ für $ \mathcal{J} = 1, \dots, 4$	98
10	Value at Risk Wert mit $\alpha = 0,995$ für $ \mathcal{J} = 1, \dots, 4$	100
11	Expected Shortfall mit $\alpha = 0,99$ für $ \mathcal{J} = 1, \dots, 4$	101
12	Value at Risk Wert mit $\alpha = 0,995$ für $ \mathcal{J} = 1, \dots, 4$	105
13	Expected Shortfall mit $\alpha = 0,99$ für $ \mathcal{J} = 1, \dots, 4$	106
14	Value at Risk Wert mit $\alpha = 0,995$ für $ \mathcal{J} = 1, \dots, 4$	108
15	Expected Shortfall mit $\alpha = 0,99$ für $ \mathcal{J} = 1, \dots, 4$	109

Tabellenverzeichnis

1	Integralschätzung klassisch und antithetische Stichprobe	49
2	Wahrscheinlichkeitsschätzung mittels Importance Sampling	53
3	Korrelationskoeffizienten von V und Z	94
4	Korrelationskoeffizienten von L und $L1$	94
5	Korrelationskoeffizienten von Z und V	99
6	Korrelationskoeffizienten von L und $L1$	99
7	Korrelationskoeffizienten von Z und V	104
8	Korrelationskoeffizienten von Z und V	107

Literatur

- [ABLM13] H. Albrecher, A. Binder, V. Loutscham, and P. Mayer. *Introduction to Quantitative Methods for Financial Markets*. Springer Basel, 2013.
- [AT02] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance* 26(7), 1487-1503, 2002.
- [BBR10] D. Bauer, D. Bergmann, and A. Reuss. Solvency ii and nested simulations - a least square monte carlo approach. [http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/3690458b48d11a78c12579ae00695536/\\$FILE/93_final%20paper_Bauer,%20et%20al.pdf](http://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/0/3690458b48d11a78c12579ae00695536/$FILE/93_final%20paper_Bauer,%20et%20al.pdf), 2010.
- [BnHH⁺16] M. Brokate, n. Henze, F. Hettlich, A. Meister, G. Schranz-Kirlinger, and Th. Sonar. *Grundwissen Mathematikstudium, Höhere analysis, Numerik und Stochastik*. Springer-Verlag Berlin Heidelberg, 2016.
- [CDS10] R. Cont, R. Deguest, and C. Scandolo. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance* 10(6), 593-606, 2010.
- [CF17] M. Cambou and D. Filipović. Replicating portfolio approach to capital calculation. *Finance and Stochastics (2018)*22, 181-203, 2017.
- [FHT96] L. Fahrmeir, A. Hamerle, and G. Tutz. *Multivariate Statistische Verfahren, 2. überarbeitete Auflage*. Walter de Gruyter Berlin - New York, 1996.
- [FS11] H. Föllmer and A. Schied. *Sochastic Finance: An Introduction in Discrete Time, 3rd Edition*. Walter de Gruyter GmbH & Co. KG, Berlin/New York, 2011.
- [GJ10] M. B. Gordy and S. Juneja. Nested simulation in portfolio risk measurement. *Management Science* 56(10), 1833-1848, 2010.
- [GKW18] R. Gillenkirch, U. Kamps, and J. Weber. Risiko. <http://wirtschaftslexikon.gabler.de/definition/risiko-44896/version-268200>, 2018.
- [Gla04] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag New York, Inc., 2004.
- [GNR12] T. Müller Gronbach, E. Novak, and K. Ritter. *Monte Carlo- Algorithmen*. Springer-Verlag Berlin Heidelberg, 2012.

- [Hub04] P. J. Huber. *Robust Statistics*. John Wiley & Sons, Inc., 2004.
- [Kle13] A. Klenke. *Wahrscheinlichkeitstheorie, 3. Auflage*. Springer-Verlag Berlin Heidelberg, 2013.
- [Kre18] J. Kremer. *Marktrisiken, Portfoliotheorie und Risikomaße*. Springer-Verlag GmbH Deutschland, 2018.
- [KW12] M. Kriele and J. Wolf. *Weltorientiertes Risikomanagement von Versicherungsunternehmen*. Springer-Verlag Berlin Heidelberg, 2012.
- [LW06] J. Lehn and H. Wegmann. *Einführung in die Statistik, 5. Auflage*. B. G. Teubner Verlag/GWV Fachverlage GmbH, Wiesbaden, 2006.
- [Nat18] J. Natolski. *Replizierende Portfolios in der Lebensversicherung, Mathematische Fundierung und Analyse*. Springer Fachmedien Wiesbaden GmbH, 2018.
- [PS15] A. Pelsser and J. Schweizer. The difference between lsmc and replicating portfolio in insurance liability modeling. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2557383, 2015.
- [Ric09] Richtlinie 2009/138/eg des europäischen parlaments und des rates. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:335:0001:0155:de:PDF>, 2009.
- [Zha09] H. Zhang. Pricing asian options using monte carlo methods. www.diva-portal.org/smash/get/diva2:301070/FULLTEXT01.pdf, 2009.