



TECHNISCHE UNIVERSITÄT WIEN

SEMINAR AUS FINANZ- UND VERSICHERUNGSMATHEMATIK

Monte-Carlo-Varianzreduktion

Autor:
Hannes HIRBER
1226611

Betreuer:
Privatdoz. Dipl-Ing. Dr.techn.
Stefan GERHOLD

28. Juli 2015

Inhaltsverzeichnis

1	Vorwort	2
2	Was ist das Monte-Carlo-Verfahren?	3
2.1	Methoden	4
2.2	Beispiel	5
3	Monte-Carlo in der Finanzmathematik	6
3.1	Komplexität	6
3.2	Amerikanische Optionen	6
4	Nachteile von Monte Carlo	7
4.1	Grundidee der Varianzreduktion	7
4.2	Vorgehensweise	8
5	Kontrollvariablen (Control Variates)	9
5.1	Idee	9
5.2	Beobachtungen	10
5.3	Beispiel	10
5.4	Multiple Variates	11
6	Antithetische Variablen	12
6.1	Idee	12
6.2	Beispiel	13
7	Geschichtete Stichproben (stratified sampling)	14
7.1	Idee	14
7.2	Beispiel	15
8	Latin Hypercube Sampling	16
8.1	Idee	16
8.2	Beispiel	17
9	Matching Underlying Assets	19
9.1	Moment matching through path adjustments	19
9.1.1	Beispiel	20
9.2	Weighted Monte Carlo	20
10	Gewichtete Stichproben (Importance Sampling)	22
10.1	Idee	22
10.2	Beispiel	23
11	Fazit	24

1 Vorwort

In meiner Seminararbeit beschäftige ich mit dem Thema der Monte-Carlo-Varianzreduktion. Dabei halte ich mich großteils an das vierte Kapitel des Buches "*Monte Carlo methods in financial Engineering*" von Paul Glasserman.

Da ich zu Beginn der Arbeit eigentlich gar keine Ahnung von diesem Thema hatte, habe ich einige Zeit gebraucht um mich einzuarbeiten. Im nachhinein bin ich froh, dass ich mich damit beschäftigt habe, da Monte-Carlo-Simulationen in der Praxis einen großen Anwendungsbereich haben und vor allem auch im Finanz- und Versicherungsbereich häufig zur Verwendung kommen. Sie sind allerdings nur zielführend, wenn die Varianz der Simulation ausreichend klein gehalten werden kann.

Einige Methoden zur Varianzreduktion sind relativ einfach nachzuvollziehen und die Durchführung scheint machbar, während andere relativ kompliziert sind und nur schwer umzusetzen sind.

Durch diese Arbeit wird zumindest überblicksmäßig eine Idee gegeben, wie die verschiedenen Methoden funktionieren.

Überblick

Für das allgemein Verständniss wird im zweiten Kapitel ein kurzer Einblick in die Idee die hinter Monte-Carlo-Methoden sowie deren Durchführung und Anwendungsbereiche in der Praxis gegeben.

Im dritten Kapitel wird dann konkret der Anwendungsbereich von Monte-Carlo-Simulationen in der Finanzmathematik erörtert, während im vierten Kapitel dann endlich das große Problem thematisiert wird: die große Varianz von Monte-Carlo-Simulationen.

Damit sind wir dann endlich beim eigentlichen Thema meiner Arbeit: Wie kann die Varianz einer solchen Simulation so klein wie möglich gehalten werden?

Glasserman stellt in seinem Buch dazu mehrere Methoden vor, die in den darauffolgenden Kapiteln vorgestellt werden.

Im letzten Kapitel folgt noch eine Analyse über die Durchführbarkeit und Sinnhaftigkeit der einzelnen Methoden im Vergleich, wobei auch der Aufwand und die gegebenen Voraussetzungen eine große Rolle spielen werden.

2 Was ist das Monte-Carlo-Verfahren?

Das Monte-Carlo-Verfahren ist eine Methode aus der Stochastik. Dabei wird versucht, mithilfe der Wahrscheinlichkeitstheorie analytisch nicht oder nur aufwändig lösbare mathematische Probleme numerisch zu lösen. Als Basis wird dabei eine große Zahl an gleichartigen Zufallsexperimenten hergenommen. Diese können entweder real durchgeführt werden, oder durch eine Computersimulation generiert werden.

Die Idee dazu stammt von Enrico Fermi, der sich damit in den 1930er Jahren auseinandergesetzt hat. Konkret ausgeführt wurden die ersten Monte-Carlo-Simulationen 1946 von Stanislaw Ulam und John von Neumann. Von Neumann benannte die Methode nach der Spielbank Monte Carlo in Monaco.

Monte-Carlo-Simulationen werden häufig verwendet, wenn:

- eine alternative Lösung zur analytischen eines rein mathematischen Problems benötigt wird:
 - Approximation der Zahl Pi mithilfe einer zufälligen "Berechnung eines Quadrates auf dem Einheitskreis
 - Berechnung des Integrals einer Funktion über dem Intervall $[0, 1]$ und auch höherdimensionale Integrale (über dem Einheitswürfel)
- Verteilungseigenschaften von Zufallsvariablen mit unbekanntem Verteilungstyp gesucht werden:
 - Ermittlung einer empirischen Verteilungsfunktion
 - Schätzung von Verteilungsparametern
- eine Nachbildung von komplexen Prozessen konstruiert werden soll, die nicht direkt analysiert werden können:
 - Produktionsprozesse in einem Fertigungsunternehmen
 - Wetter und Klima der Erde
 - Rekonstruktionsverfahren in der Nuklearmedizin

Auch Probleme mit statistischem Verhalten kann man gut mit dem Monte-Carlo-Verfahren simulieren, was vor allem in der Physik eine große Rolle spielt.

Mathematisch gesehen ist das System ein Wahrscheinlichkeitsgewichteter Weg im Phasenraum Ω . Somit ist es besonders gut dazu geeignet, statistische Mittelwerte einer Größe \mathcal{A} mithilfe von normierten statistischen Gewichten $\mathcal{P}(x)$ zu berechnen:

$$\langle \mathcal{A} \rangle = \sum_{x \in \Omega} \mathcal{P}(x) \mathcal{A}(x)$$

Wenn der Raum Ω so groß ist, dass die Summation nicht durchgeführt werden kann, erzeugt man stattdessen eine Markov-Kette x_1, x_2, x_3, \dots mit Zuständen in Ω , wobei die Häufigkeit eines Zustandes wie das vorgegebene Gewicht $\mathcal{P}(x)$ verteilt ist. Damit lässt sich der Erwartungswert als arithmetisches Mittel aus den Zuständen der Markov-Kette berechnen:

$$\langle \mathcal{A} \rangle = \frac{1}{N} \sum_{i=1}^N \mathcal{A}(x_i)$$

Dieses Ergebniss basiert auf dem Gesetz der großen Zahlen. Wichtig ist, dass die Markov-Kette den gesamten Raum abdecken muss, also ergodisch sein muss. So eine Markov-Kette ist oft sehr schwer zu finden.

2.1 Methoden

Es gibt nun verschiedene Methoden, eine Monte-Carlo-Simulation zu konstruieren:

- *Metropolis-Monte-Carlo*
Hier wird nach einem von Nicholas Metropolis publizierten Algorithmus vorgegangen, der sich von der Monte-Carlo-Integration ableitet und der zur Untersuchung statistisch-mechanischer Systeme dient.
- *Sequentielle Monte-Carlo-Methode*
Es wird versucht den Systemzustand als Funktion der Zeit auf Basis einer Reihe von Beobachtungen des Systems und A-priori-Kenntnissen der Systemdynamik zu schätzen. Dabei wird die Wahrscheinlichkeitsdichte des Zustandes diskret durch eine Menge von Partikeln angenähert.
- *Quanten-Monte-Carlo-Methoden*
Sie werden zur Berechnung von Physikalischen Observablen in quantenfeldtheoretischen Modellen benutzt. (z.B. in der Festkörperphysik: Hubbard-Modell, tJ-Modell)
- *kinetische Monte-Carlo-Methode*
Diese erlaubt es den zeitlichen Fortschritt eines Systems zu simulieren.
- *MCE, MCS und MCR Methode*
Hier werden normierte Wertermittlungsverfahren (Monte-Carlo-Ertragswert, Monte-Carlo-Sachwert und Monte-Carlo-Residualwertmethode) in die Simulation miteinbezogen.

2.2 Beispiel

Auf der Abbildung sehen wir, wie bei der Approximation der Zahl Pi durch eine Monte-Carlo-Simulation vorgegangen wird:

Zuerst haben wir ein Quadrat mit der Breite 1 gegeben. In diese wird ein Viertel des Einheitskreises eingeschrieben. Nun werden zufällig Punkte auf dem Quadrat verteilt. Diese werden anschließend gezählt und es wird das Verhältnis der Punkte innerhalb des Kreises und jener außerhalb aufgestellt.

Je mehr Punkte verwendet werden, umso besser wird die Zahl $\frac{\pi}{4}$ approximiert.

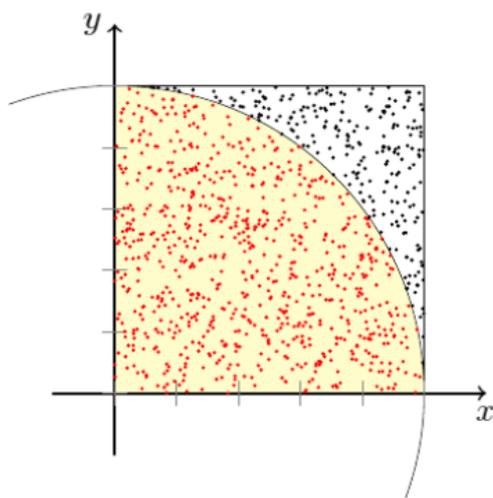


Abbildung 1: Approximation von Pi

3 Monte-Carlo in der Finanzmathematik

Monte-Carlo-Simulationen (MCS) werden in der Finanzmathematik dazu verwendet Portfolios, Kapitalanlagen und andere Finanzinstrumente zu bewerten und zu analysieren, indem die verschiedenen Risikofaktoren simuliert werden und danach der Mittelwert der Ergebnisse bestimmt wird.

Sie wurden erstmals von 1964 von David Hertz in die Finanzmathematik eingeführt, als er sich mit ihrer Anwendung in der Unternehmensfinanzierung (Corporate Finance) auseinandersetzte und dazu auch ein Paper veröffentlichte.

1977 verwendete Phelim Boyle MCS erstmals zur Bewertung von Derivaten. Dabei werden mehrere tausend mögliche (aber zufällige) Preisprozesse simuliert, die den Payoff der Option wiedergeben. Diese Payoffs werden dann gemittelt und auf den momentanen Zeitpunkt diskontiert, was dann den momentanen Wert der Option ergibt.

Außerdem werden MCS auch in der Portfoliobewertung und in der persönlichen Finanzplanung verwendet, indem der gesamte Markt und nicht nur die einzelne Option simuliert wird.

3.1 Komplexität

Meist wird ein spezifisches Integral (z.B. der arbitragefreie Preis eines Derivats) gesucht. Diese können oft einfach analytisch berechnet, oder mithilfe numerischer Integration oder partiellen Differentialgleichungen bewertet werden. Sobald aber mehr als drei oder vier Zustandsvariable vorkommen, gibt es keine Formeln wie die Black-Scholes Formeln mehr und auch numerische Methoden, wie z.B. die Annäherung durch das Binomialmodell, werden sehr aufwändig. Hier konvergieren Monte-Carlo-Simulationen schneller als die numerische Integration und kommen deshalb oft zur Anwendung.

Bei einfacheren Problemen kommt die MCS hingegen nicht zur Anwendung, da sie sehr zeitaufwändig und rechenintensiv sind.

3.2 Amerikanische Optionen

Bei Amerikanischen Optionen sind Monte-Carlo-Methoden schwieriger zu verwenden, da sie im Gegensatz zu partiellen Differentialgleichungen nur den Wert der Option zu einem gegebenen Zeitpunkt schätzen. Allerdings werden auch die Optionswerte zu den Zeitpunkten zwischen der Startzeit und dem Ablaufdatum benötigt, um bereits frühzeitig zu handeln. Diese werden beim Black-Scholes Modell einfach berechnet, da hier die Berechnung rückwärts verläuft. Bei einer Monte-Carlo Simulation sind diese schwieriger zu erhalten, aber sie können zum Beispiel mithilfe des "Methode der kleinsten Quadrate"-Algorithmus von Carrere berechnet werden.

4 Nachteile von Monte Carlo

Die Monte-Carlo-Simulationen bieten zwar eine gute Möglichkeit, ein schwieriges Problem relativ einfach zu lösen, allerdings birgt diese Art des Lösungsansatzes auf gewisse Gefahren:

- Eine Monte-Carlo-Simulation konvergiert nur sehr langsam, weshalb es eine große Zahl an simulierten Werten benötigt wird, um ein genaues Resultat zu erhalten.
- Der Schätzer für den Preis eines Assets ist nur eine Zufallsvariable. Diese Unsicherheit kann zu schlechten Entscheidungen im Risikomanagement führen.

Es ist deshalb wichtig, die Varianz so klein wie möglich zu halten. Es gibt dazu verschiedene vorgehensweisen, von denen wir einige später noch kennenlernen werden, allerdings folgen alle einem bestimmten Grundprinzip.

4.1 Grundidee der Varianzreduktion

Bei Monte-Carlo-Simulationen soll ein spezifischer Wert s (zum Beispiel ein Integral) durch einen Erwartungswert ausgedrückt werden:

$$s = E[f(X)]$$

Ist X_1, \dots, X_n eine Stichprobe von unabhängigen Zufallsvariablen, mit derselben Verteilung wie X , so kann s mit großen n durch arithmetisches Mittel angenähert werden:

$$S_n = \frac{1}{n} \sum_{i=1}^n f(X_i)$$

Die obige Gleichung ist korrekt, da wegen dem Gesetz der Großen Zahlen und der Linearität des Erwartungswertes gilt:

$$E[S_n] = s$$

Die Genauigkeit der Schätzung lässt sich nun mit der Varianz von S_n messen. Wegen der Unabhängigkeit der X_i gilt:

$$\text{Var}(S_n) = \frac{1}{n} \text{Var}(f(X))$$

Daraus resultiert eine Konvergenzordnung von $\mathcal{O}(\frac{1}{\sqrt{n}})$ der Standardabweichung von S_n . Im Allgemeinen lässt sich das Konvergenzverhalten von S_n nicht verbessern, weshalb

man bei $Var(f(X))$ ansetzen muss. Man versucht die Funktion f und die Verteilung von X so zu wählen, dass die Varianz möglichst klein wird. Oft ist die Varianz von $f(X)$ aber nicht bekannt und muss durch die Stichprobenvarianz

$$\frac{1}{n-1} \sum_{i=1}^n (f(X_i) - S_n)^2$$

geschätzt werden.

4.2 Vorgehensweise

Es gibt 2 umfassende Strategien, die zu einer Varianzreduktion führen:

- 1) Ausnutzen von lenkbaren Eigenschaften eines Modells um den Output zu verbessern
- 2) Vermindern der Variabilität des Simulations-Inputs

Im vierten Kapitel des Buches *Monte Carlo Methods in Financial Engineering* von Paul Glasserman werden unter anderem folgende verschiedene Methoden der Varianzreduktion diskutiert, die ich anschließend kurz vorstellen werde:

- Kontrollvariablen (Control Variates)
- Antithetische Variablen (antithetic variates)
- Geschichtete Zufallsstichproben (stratified sampling)
- Latin Hypercube Sampling
- Matching Underlying Assets
- Gewichtete Stichproben (Importance Sampling)

Generell geht es darum, gewisse Eigenschaften eines spezifischen Problems auszunutzen, anstatt dies bei der allgemeinen Methode zu versuchen.

5 Kontrollvariablen (Control Variates)

5.1 Idee

Bei dieser Methode wird versucht, die Information über den Fehler eines Schätzers einer bekannten Menge auszunutzen, um den Fehler einer unbekannt Menge zu reduzieren. Dazu wird zuerst die Simulation durchgeführt, allerdings wird an die unbekannt Variable, die geschätzt werden soll, eine Variable gekoppelt, deren wirklicher Wert bereits bekannt ist. Es soll also $E[Y_i]$ geschätzt werden: Y_1, \dots, Y_n seien die Ergebnisse einer Simulation. Falls diese unabhängig und gleichverteilt sind, so ist der übliche Schätzer durch

$$\bar{Y} = \frac{(Y_1 + \dots + Y_n)}{n}$$

gegeben.

Dieser Schätzer ist unverzerrt und konvergiert für $n \rightarrow \infty$ laut dem Gesetz der großen Zahlen mit Wahrscheinlichkeit 1 gegen den richtigen Wert.

Nun wird zu jedem Ergebnis Y_i auch ein X_i simuliert, wobei die Paare $(X_i, Y_i), i = 1, \dots, n$ unabhängig und gleichverteilt sind und $E[X]$ bekannt sei.

Dann können wir für jedes fixe b

$$Y_i(b) = Y_i - b(X_i - E[X])$$

berechnen.

Mit dem Durchschnitt

$$\bar{Y}_i(b) = \bar{Y} - b(\bar{X} - E[X]) = \frac{1}{n} \sum_{i=1}^n (Y_i - b(X_i - E[X]))$$

erhält man nun einen Kontrollvariablen-Schätzer, wobei der beobachtete Fehler $\bar{X} - E[X]$ als Kontrolle dient.

Auch dieser Schätzer ist unverzerrt und konsistent, wie wir nun zeigen werden:

Die Varianz ist gegeben durch:

$$Var[Y_i(b)] = \sigma_Y^2 - 2b\sigma_X\sigma_Y\rho_{XY} + b^2\sigma_X^2$$

wobei $\sigma_Y^2 = Var[Y]$, $\sigma_X^2 = Var[X]$ und ρ_{XY} die Korrelation zwischen X und Y darstellt.

Der optimale Parameter b , der die Varianz minimiert, ist nun gegeben durch

$$b^* = \frac{\sigma_Y}{\sigma_X} \rho_{XY} = \frac{Cov[X, Y]}{Var[X]}$$

Durch Substituieren und Vereinfachen kommt man darauf, dass das optimale Verhältnis zwischen kontrolliertem und unkontrolliertem Schätzer genau $1 - \rho_{XY}^2$ ist.

5.2 Beobachtungen

Die Effizienz einer Kontrollvariablen wird also durch die Korrelation zwischen der gesuchten Variablen und der Kontrollvariablen bestimmt.

Falls der Aufwand für die Replikation mit und ohne der Kontrollvariante ungefähr derselbe bleibt, so gibt $1 - \rho_{XY}^2$ die Beschleunigung der Rechenintensität durch deren Benutzung an.

Die Anzahl von Replikationen von Y_i die benötigt wird, um dieselbe Varianz wie unter Verwendung der Kontrollvariante zu erhalten, beträgt also:

$$\frac{n}{1 - \rho_{XY}^2}$$

Da in der Praxis σ_Y und ρ_{XY} meist nicht bekannt sind, muss oft ein Schätzer für b^* verwendet werden. Durch das Gesetz der großen Zahlen konvergiert dieser zwar gegen b^* , aber trotzdem kann es zu einer gewissen Verzerrtheit des Schätzers kommen.

5.3 Beispiel

Um uns die Anwendung und Bedeutung dieser Methoden einmal vor Augen zu führen wollen wir uns nun mit einem spezifischen Beispiel auseinandersetzen.

Wir wollen das Integral

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

mithilfe von Monte-Carlo berechnen.

Dieses Integral ist der Erwartungswert von $f(U)$, wobei

$$f(x) = \frac{1}{1+x}$$

und U gleichverteilt auf $[0,1]$ sind.

Nehmen wir nun eine Stichprobe u_1, \dots, u_n her, so ist der Schätzer gegeben durch

$$I \approx \frac{1}{n} \sum_i f(u_i)$$

Wir verwenden nun $g(x) = 1 + x$ als Kontrollvariable, deren Erwartungswert durch

$$E[g(U)] = \int_0^1 (1+x) dx = \frac{3}{2}$$

gegeben ist.

Wir kombinieren die beiden, wodurch wir nun

$$I \approx \frac{1}{n} \sum_i f(u_i) + c \left(\frac{1}{n} \sum_i g(u_i) - \frac{3}{2} \right)$$

erhalten.

Für $n = 1500$ und den geschätzten optimalen Koeffizienten $b^* \approx 0,4773$ erhalten wir:

Method	Schätzwert	Varianz
Klassische Schätzung	0.69475	0.01947
Mit Kontrollvariablen	0.69295	0.00060

Natürlich können wir dieses Integral auch analytisch berechnen. Der exakte Wert ist

$$I = \ln(2) \approx 0.693147$$

5.4 Multiple Variates

Anstatt nur einer Kontrollvariablen können auch mehrere verwendet werden, die Formeln müssen dann nur entsprechend angepasst werden.

Es sein nun

$$\begin{pmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma_{XY}^T & \sigma_Y^2 \end{pmatrix}$$

die Kovarianzmatrix von den unabhängig und gleichverteilten Paaren $(X_i, Y_i), i = 1, \dots, n$. Für die Varianz ergibt sich:

$$\text{Var}[Y_i - b^T(X_i - E[X])] = \sigma_Y^2 - 2b\Sigma_{XY} + b^T\Sigma_{XX}bb^T$$

Diese wird minimiert bei:

$$b^* = \Sigma_X^{-1}\Sigma_{XY}$$

Die bisher betrachteten Methoden waren linear, es gibt aber auch nichtlineare Methoden der Kontrollvariablen. Diese sind dann der Form

$$h(E[X], y) = y$$

Beispiele dafür sind z.B.

$$\bar{Y} \frac{E[X]}{\bar{X}}$$

oder

$$\bar{Y} \exp(\bar{X} - E[X])$$

Beide sind Spezialfälle von Schätzern der Form $h(\bar{X}, \bar{Y})$ für Funktionen, die die obige Gleichung erfüllen.

6 Antithetische Variablen

6.1 Idee

Sei U eine uniform verteilte Zufallsvariable auf $[0,1]$. Die Idee der antithetischen Variablen liegt darin, für einen Weg U_1, \dots, U_n seinen antithetischen Weg $1 - U_1, \dots, 1 - U_n$ herzunehmen und anhand dieser beiden Wege eine bessere Aussage treffen zu können, ohne einen viel größeren Rechenaufwand für die Simulation zu haben. Die beiden Pfade bilden ein antithetisches Paar.

$F^{-1}(U)$ und $F^{-1}(1 - U)$ haben klarerweise dieselbe Verteilung.

Man erhält nun anhand der Implementierung der antithetischen Paare die Beobachtungen $(Y_i$ und $\tilde{Y}_i), i = 1, \dots, n$. Diese sind unabhängig und gleichverteilt.

Y_i und \tilde{Y}_i hingegen haben dieselbe Verteilung, sind aber nicht unabhängig voneinander. Der antithetische Schätzer ist der Mittelwert der $2n$ Werte:

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

Die Varianz ist gegeben durch:

$$\sigma_{AV}^2 = \text{Var} \left[\frac{Y_i + \tilde{Y}_i}{2} \right]$$

Die Antithetische-Methode reduziert die Varianz genau dann, wenn

$$\text{Var}[Y_i + \tilde{Y}_i] < 2\text{Var}[Y_i]$$

Aufgrund der gleichen Verteilung gilt:

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

Das heißt also, dass die Methode genau dann Varianz reduziert, wenn

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

Wenn f eine lineare Funktion ist, so gilt offenbar:

$$\frac{U + (1 - U)}{2} = \frac{1}{2}$$

und

$$\frac{Z + (-Z)}{2} = 0$$

Daraus folgt, dass die Antithetische-Variablen-Methode vor allem bei annähernd linearen Funktionen sehr effektiv ist.

6.2 Beispiel

Wir wollen uns auch diese Methode nun anhand eines einfachen Beispiels veranschaulichen:

Wir betrachten erneut das Beispiel

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

Mit der Monte-Carlo Methode erhalten wir bei $n=1500$ für $2n$ Einträge:

Method	Schätzwert	Varianz
Klassische Schätzung	0.69365	0.02005
Antithetische Variable	0.69399	0.00063

Die Varianz ist also tatsächlich signifikant kleiner.

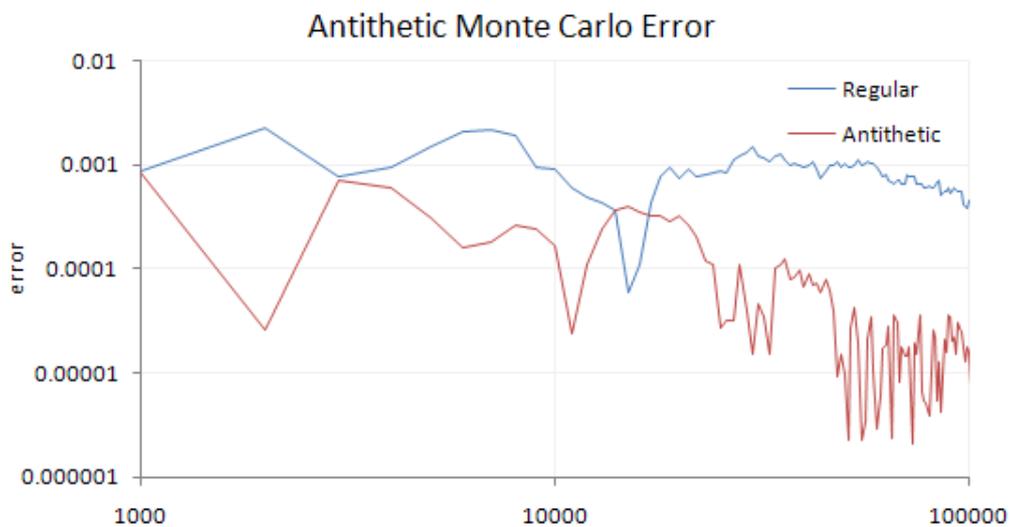


Abbildung 2: Vergleich der regulären und antithetischen Methode (bei obigem Beispiel)

7 Geschichtete Stichproben (stratified sampling)

7.1 Idee

Die Idee der Geschichteten Stichprobe liegt darin, die Grundgesamtheit in mehrere Gruppierungen, den sogenannten Schichten, zu unterteilen. Danach wird aus jeder Schicht eine Stichprobe gezogen.

Die Stichproben werden entsprechend nach dem Umfang ihrer Schichten, die bekannt sein müssen, gewichtet.

Durch die Schichtung erhält man bei einer günstigen Auswahl genauere, bei einer schlechten Auswahl aber mindestens genauso gute Ergebnisse, wie bei einer einfachen Stichprobe.

Der Erwartungswert einer Geschichteten Stichprobe ergibt sich aus:

$$E[Y] = \sum_{i=1}^K P(Y \in A_i) E[Y|Y \in A_i] = \sum_{i=1}^K p_i E[Y|Y \in A_i]$$

mit

$$p_i = P(Y \in A_i)$$

wobei K die Anzahl der Schichten ist.

Daraus folgt nun, dass

$$\hat{Y} = \sum_{i=1}^K p_i \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij} = \frac{1}{n} \sum_{i=1}^K \sum_{j=1}^{n_i} Y_{ij}$$

ein unverzerrter Schätzer ist.

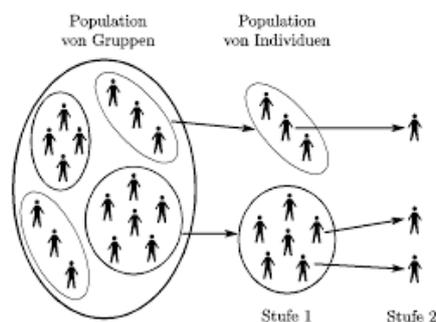


Abbildung 3: Schichten und Stichproben

7.2 Beispiel

Wir gehen von einer Gleichverteilung aus und unterteilen das Intervall $(0, 1)$ in n Schichten:

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

Jedes dieser Intervalle hat die Wahrscheinlichkeit $1/n$, U_1, \dots, U_n wobei die U_i 's unabhängig und gleichverteilt zwischen 0 und 1 sind. Wir setzen zudem

$$V_i = \frac{i-1}{n} + \frac{U_i}{n}, \quad i = 1, \dots, n$$

sodass die V_i 's gleichverteilt auf dem Intervall $[\frac{i-1}{n}, \frac{i}{n}]$ sind.

Sei nun $Y = f(U)$, also $E[Y]$ das Integral von f über $[0, 1]$. Dann ist der geschichtete Schätzer gegeben durch:

$$\hat{Y} = \frac{1}{n} \sum_{i=1}^n f(V_i)$$

Durch die Schichtung ist \hat{Y} unverzerrt.

8 Latin Hypercube Sampling

8.1 Idee

Diese Methode ist eine Erweiterung von geschichteten Stichproben in höheren Dimensionen. Es wird nun versucht, Stichproben von dem d -dimensionalen "Würfel" $[0, 1)^d$ zu nehmen. Dazu wird jede Koordinate in K Schichten eingeteilt, wodurch K^d Schichten zustandekommen. Das bedeutet, dass mindestens K^d Stichproben genommen werden müssen, damit von jeder Schicht eine genommen wurde.

Das Prinzip der Lateinischen-Würfel-Stichprobe betrachtet alle Schichten als gleichwertig und nimmt genau eine Stichprobe von jeder Schicht einer Dimension. Es verhindert somit das exponentielle Wachstum der vollständig geschichteten Stichproben, indem es nur die eindimensionalen Ränder schichtet.

Die von McKay, Conover und Beckman entwickelte und von Stein weiter analysierte Methode wird am einfachsten anhand einer Gleichverteilung über den Einheitswürfel erklärt.

Es werden die Dimension d und die Stichprobengröße K festgelegt. Für jede Koordinate $i = 1, \dots, d$ wird unabhängig eine Stichprobe $V_i^{(1)}, \dots, V_i^{(K)}$ auf dem Einheitsintervall genommen. Jedes $V_i^{(j)}$ ist gleichverteilt auf $[\frac{j-1}{K}, \frac{j}{K})$.

Wenn wir die d geschichteten Stichproben in Spalten anordnen, so erhalten wir:

$$\begin{pmatrix} V_1^{(1)} & V_2^{(1)} & \dots & V_d^{(1)} \\ V_1^{(2)} & V_2^{(2)} & \dots & V_d^{(2)} \\ V_1^{(3)} & V_2^{(3)} & \dots & V_d^{(3)} \\ \vdots & \vdots & & \vdots \\ V_1^{(K)} & V_2^{(K)} & \dots & V_d^{(K)} \end{pmatrix}$$

Jede Zeile gibt Koordinaten eines Punktes in $[0, 1)^d$ an, wobei die erste Zeile einen Punkt in $[0, \frac{1}{K})^d$, die zweite Zeile einen Punkt in $[\frac{1}{K}, \frac{2}{K})^d$ usw. angibt. Die Punkte liegen also in der Diagonale des Einheitswürfels. Nun seien π_1, \dots, π_d zufällige Permutationen von $1, \dots, K$. $\pi_j(i)$ gibt den Wert von i bezüglich der j -ten Permutation an. Die Zeilen der Matrix

$$\begin{pmatrix} V_1^{\pi_1(1)} & V_2^{\pi_2(1)} & \dots & V_d^{\pi_d(1)} \\ V_1^{\pi_1(2)} & V_2^{\pi_2(2)} & \dots & V_d^{\pi_d(2)} \\ V_1^{\pi_1(3)} & V_2^{\pi_2(3)} & \dots & V_d^{\pi_d(3)} \\ \vdots & \vdots & & \vdots \\ V_1^{\pi_1(K)} & V_2^{\pi_2(K)} & \dots & V_d^{\pi_d(K)} \end{pmatrix}$$

geben auch weiterhin Punkte in $[0, 1)^d$ an, allerdings befinden sich diese nun nicht mehr in der Diagonale, sondern sind gleichverteilt über dem Einheitswürfel.

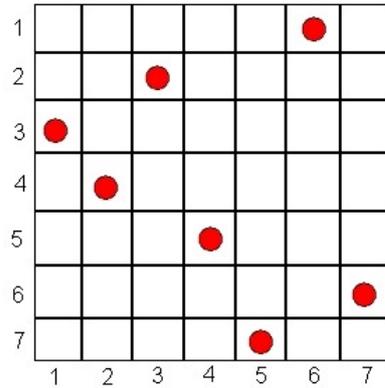


Abbildung 4: Latin Hypercube Sample im 2-dimensionalen.

In der Abbildung sehen wir ein Beispiel für den zweidimensionalen Fall, wobei hier $K = 7$ gilt.

8.2 Beispiel

Um nun die Effizienz dieser Methode zu untersuchen, betrachten wir wieder den Einheitswürfel $[0, 1]^d$. Wir wollen nun

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

für eine quadratisch integrierbare Funktion $f : [0, 1]^d \rightarrow \mathcal{R}$ schätzen.

Der Standard Monte-Carlo Schätzer des Integrals kann als

$$\bar{\alpha}_f = \frac{1}{K} \sum_{j=0}^{K-1} f(U_{jd+1}, U_{jd+2}, \dots, U_{jd+d})$$

mit U_1, U_2, \dots unabhängige Gleichverteilungen.

Die Varianz dieses Schätzers ist durch σ^2/K gegeben, wobei $\sigma^2 = \text{Var}[f(U_1, \dots, U_d)]$.

Wir setzen nun:

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

und definieren den Schätzer:

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

McKay gelang es zu zeigen, dass die Varianz dann folgendermaßen aussieht:

$$\text{Var}[\hat{\alpha}_f] = \frac{\sigma^2}{K} + \frac{K-1}{K} \text{Cov}[f(V^{(1)}), f(V^{(2)})]$$

Falls die Funktion f monoton in jeder Koordinate ist, dann gilt für jedes quadratisch-integrierbare f und $K \geq 2$ gilt:

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

Die Varianz einer "Lateinischen-Würfel-Stichprobe" der Größe K ist also auf keinen Fall größer als die Varianz einer unabhängig gleichverteilten Stichprobe der Größe $K-1$.

9 Matching Underlying Assets

In diesem Kapitel werden Methoden diskutiert, die sich mit dem Problem auseinandersetzen, dass gewisse Stichprobenmittel mit den Werten der Grundgesamtheit übereinstimmen. Vor allem in der Finanzmathematik spielen diese Methoden eine wichtige Rolle, da bei einer falschen Bepreisung von Assets eine Arbitragemöglichkeit entstehen könnte. Wir betrachten zwei verschiedene Methoden: *moment matching*, das auf Transformationen von simulierten Pfaden basiert, und eine Methode, die den Pfaden verschiedene Gewichte zuweist um die Momente anzupassen. Im Vergleich zu Kontrollvariablen oder auch im Vergleich untereinander ergeben sich für kleine Stichproben große Unterschiede, während die Methoden bei größeren Zahlen äquivalent werden.

9.1 Moment matching through path adjustments

Die Idee Pfade so zu transformieren, dass Momente angepasst werden, wird am einfachsten in einem Model mit einem risikoneutralen Maß und einer konstanten Zinsrate r erklärt, in dem nur ein einzelnes zugrundeliegendes Asset (underlying asset) $S(t)$ simuliert wird. Wenn das Asset keine Dividenden auszahlt, so wissen wir, dass $E[S(t)] = e^{rt}S(0)$ gilt. Wir simulieren nun n unabhängige Kopien S_1, \dots, S_n des Prozesses und definieren das Stichprobenmittel als

$$\bar{S}(t) = \frac{1}{n} \sum_{i=1}^n S_i(t).$$

Für endliche n wird das Stichprobenmittel nicht immer mit $E[S(t)]$ übereinstimmen, die Simulation bepreist das zugrundeliegende Asset falsch, sodass

$$e^{-rt}\bar{S}(t) \neq S(0),$$

wobei die rechte Seite den aktuellen Preis des Assets und die linke Seite deren Schätzer durch die Simulation darstellt.

Eine mögliche Abhilfe ist durch eine Transformation der simulierten Pfade gegeben, indem

$$\tilde{S}_i(t) = S_i(t) \frac{E[S(t)]}{\bar{S}(t)}, i = 1, \dots, n,$$

oder

$$\tilde{S}_i(t) = S_i(t) + E[S(t)] - \bar{S}(t), i = 1, \dots, n$$

gesetzt wird und dann $\tilde{S}_i(t)$ anstelle von S_i verwendet wird, um die Derivate zu bepreisen.

Beide garantieren, dass das Stichprobenmittel mit $E[S(t)]$ übereinstimmt. Diese Methode wird "moment matching" genannt.

Die Methode hat zur Folge, dass es zur sogenannten Put-Call-Parität kommt, das heißt es gilt

$$(a - b)^+ - (b - a)^+ = a - b$$

was folgende Bedingung impliziert:

$$e^{-rT} E[(S(T) - K)^+] - e^{-rT} E[(K - S(T))^+] = S(0) - e^{-rT} K$$

Beide Transformationen verändern den simulierten Prozess und verursachen üblicherweise eine Verzerrung von Schätzungen die aus den angepassten Pfaden berechnet wurden. Diese Verzerrung verschwindet mit größer werdender Stichprobengröße n und ist normalerweise $\mathcal{O}(1/n)$.

9.1.1 Beispiel

Für unabhängige gleichverteilte normale Zufallsvektoren ist es äquivalent, ob man über das Stichprobenmittel zentriert, oder bedingt auf Stichprobenmittel als Mittelwert der Grundgesamtheit.

Um das zu sehen, seien X_1, \dots, X_n unabhängig nach $N(\mu, \Sigma)$ Zufallsvektoren. Die angepassten Vektoren

$$\tilde{X} = X_i - \bar{X} + \mu$$

haben den Mittelwert μ . Zudem sind sie untereinander normal mit

$$\begin{pmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_n \end{pmatrix} \sim N \left(\begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{pmatrix} (n-1)\Sigma/n & -\Sigma/n & \cdots & -\Sigma/n \\ -\Sigma/n & (n-1)\Sigma/n & & \vdots \\ \vdots & & \ddots & -\Sigma/n \\ -\Sigma/n & & -\Sigma/n & (n-1)\Sigma/n \end{pmatrix} \right)$$

Dies kann recht einfach mit Lineartransformationen gezeigt werden. Allerdings ist dies auch die gemeinsame Verteilung von X_1, \dots, X_n mit $\bar{X} = \mu$.

9.2 Weighted Monte Carlo

Eine andere Möglichkeit der Anpassung von zugrundeliegenden Preisen bei endlichen Stichproben liegt darin, den verschiedenen Pfaden ein Gewicht zu geben.

Wir gehen wieder davon aus, dass die Simulation das zugrundeliegende Asset wieder falsch bepreist, also dass das Stichprobenmittel $\bar{S}(t)$ ungleich $e^{rt}S(0)$ ist.

Nun verändern wir aber nicht die simulierten Werte $S_1(t), \dots, S_n(t)$, sondern wir wählen

Gewichte $\omega_1, \dots, \omega_n$, welche die Gleichung

$$\sum_{i=1}^n \omega_i S_i(t) = e^{rt} S(0)$$

erfüllen. Diese Gewichte verwenden wir dann, um den erwarteten Payoff einer Option zu berechnen. Zum Beispiel ergibt sich die Schätzung

$$e^{-rt} \sum_{i=1}^n \omega_i (S_i(t) - K)^+$$

für den Preis eines Call mit Maturität K .

Die Methode kann noch verallgemeinert werden. Wollen wir etwa $E[Y]$ schätzen und wissen den Mittelvektor $\mu_X = E[X]$ für einen d -wertigen Zufallsvektor. X kann beispielsweise die Preise von zugrundeliegenden Assets an zukünftigen Zeitpunkten, die Stärke dieser Preise oder die diskontierten Payoffs von lenkbaren Optionen angeben. Wir simulieren nun unabhängige gleichverteilte Replikationen $(X_i, Y_i), i = 1, \dots, n$ des Paares (X, Y) und suchen nun Gewichte $\omega_i, i = 1, \dots, n$, die

$$\sum_{i=1}^n \omega_i X_i = \mu_X$$

erfüllen und verwenden diese, um $E[Y]$ zu berechnen. Dieser ergibt sich aus:

$$\sum_{i=1}^n \omega_i Y_i.$$

Außerdem wollen wir, dass

$$\sum_{i=1}^n \omega_i = 1$$

gilt. Üblicherweise ist die Anzahl der Bedingungen d kleiner als die Anzahl der Replikationen n , sodass die Gewichte nicht eindeutig bestimmt sind. Wir wählen unsere Gewichte, indem wir eine Funktion $H : \mathcal{R}^d \rightarrow \mathcal{R}$ wählen und das Optimierungsproblem

$$\min H(\omega_1, \dots, \omega_n)$$

lösen.

10 Gewichtete Stichproben (Importance Sampling)

10.1 Idee

Bei der gewichteten Stichprobe wird versucht, wichtigen Ergebnissen mehr Gewicht zu verleihen als anderen, um die Effizienz der Stichprobe zu erhöhen. Dabei gehen wir davon aus, dass das Problem folgende Form hat:

$$\alpha = E[h(X)] = \int h(x)f(x)dx$$

wobei f die Wahrscheinlichkeitsdichte, X eine Zufallsvariable aus \mathbb{R}^d und h eine Funktion von \mathbb{R}^d auf \mathbb{R} sei.

Der übliche Monte-Carlo Schätzer ist durch

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n h(X_i)$$

gegeben, wobei $X_i, i = 1, \dots, n$ unabhängig ist.

Sei nun g eine weitere Wahrscheinlichkeitsdichte auf \mathbb{R}^d , für die gelten soll:

$$f(x) > 0 \Rightarrow g(x) > 0$$

Dadurch können wir α alternativ darstellen als

$$\alpha = \int h(x) \frac{f(x)}{g(x)} g(x) dx$$

Dieses Integral kann als Erwartungswert bezüglich der Dichte g interpretiert werden:

$$\alpha = \tilde{E} \left[h(X) \frac{f(X)}{g(X)} \right]$$

wobei nun X nach g verteilt ist.

Der mit g verbundene Importance-Sampling-Schätzer ist gegeben durch

$$\hat{\alpha}_g = \frac{1}{n} \sum_{i=1}^n h(X_i) \frac{f(X_i)}{g(X_i)}$$

wobei

$$\frac{f(X_i)}{g(X_i)}$$

die Radon-Nikodym-Ableitung bei X_i ist.

Es gilt außerdem

$$\tilde{E}[\hat{\alpha}_g] = \alpha$$

Also ist $\hat{\alpha}_g$ ein unverzerrter Schätzer für α .

Um die Varianzen mit und ohne Importance Sampling zu vergleichen, reicht es also, die zweiten Momente zu vergleichen.

Der Erfolg dieser Variante hängt sehr von der Wahl von g ab, bei einer schlechten Wahl kann auch ein sehr viel schlechteres Ergebnis als bei der gewöhnlichen Methode erzielt werden.

10.2 Beispiel

Auch hier wollen wir uns wieder ein Beispiel zum allgemeinen Verständniss anschauen. Unser Ziel ist es, den Wechsel des Mittelwertes einer Normalverteilung bei einem Maßwechsel zu berechnen:

Sei f die eindimensionale Standardnormalverteilung und g die eindimensionale Normalverteilung mit Mittelwert μ und Varianz 1.

Dann kann einfach gezeigt werden, dass folgende Gleichung gilt:

$$\prod_{i=1}^m \frac{f(Z_i)}{g(Z_i)} = \exp\left(-\mu \sum_{i=1}^m Z_i + \frac{m}{2}\mu^2\right)$$

Wenn wir allgemeiner g_i mit Mittelwert μ_i festlegen, so gilt:

$$\prod_{i=1}^m \frac{f(Z_i)}{g_i(Z_i)} = \exp\left(-\sum_{i=1}^m \mu_i Z_i + \frac{1}{2} \sum_{i=1}^m \mu_i^2\right)$$

Wenn wir eine Brown'sche Bewegung auf dem Gitter $0 = t_0 < \dots < t_m$ simulieren, indem wir

$$W(t_n) = \sum_{i=1}^n \sqrt{t_i - t_{i-1}} Z_i$$

setzen, so ist obige Gleichung das Wahrscheinlichkeitsverhältnis eines Maßwechsels, das den Mittelwert $\mu\sqrt{t_i - t_{i-1}}$ zum Brown'schen Inkrement über $[t_{i-1}, t_i]$ addiert.

11 Fazit

Zum Schluss wollen wir uns noch einmal kurz die verschiedenen Möglichkeiten zu einer Varianzreduktion anschauen und kurz diskutieren, welche Methoden praktikabel sind und wann welche Methode am wahrscheinlichsten zu einer befriedigenden Lösung führt. Das Problem liegt darin, dass es sehr schwierig ist herauszufinden, welche Methode man verwenden sollte. Wichtige Faktoren bei der Wahl sind die vorhandenen Informationen und die verfügbare Zeit.

Man kann die einzelnen Methoden nach Aufwand und Effizienz bewerten:

- Antithetische Stichproben benötigen keine spezifische Information und sind einfach zu implementieren. Es bringt aber auch kaum bzw. nur selten große Varianzreduktion
- Kontrollvariable sind ebenfalls recht einfach zu implementieren, zudem ist gewährleistet, dass die Varianz nicht größer wird.
- Geschichtete Stichproben sind etwas komplizierter, da nicht nur der Mittelwert, sondern die Verteilung einer Variablen benötigt wird. Auch hier wird garantiert, dass die Varianz nicht erhöht wird
- Latin Hypercube Sampling ist eine Verallgemeinerung der geschichteten Stichprobe, womit es generell ein komplexeres System ist. In gewissen Fällen kann es allerdings auch ziemlich einfach zu implementieren sein.
- Bei gewichteten Stichproben muss die Verteilung von der neuen Dichtefunktion sehr sorgfältig gewählt werden, da es sonst sogar zu unendlicher Varianz kommen kann. Wenn sie gut gewählt wird, können mit ihr aber auch sehr gute Ergebnisse zustandekommen.

Quellen

- Monte Carlo Methods in Financial Engineering
- https://en.wikipedia.org/wiki/Monte_Carlo_methods_in_finance Zugriff 15.07.2015
- <https://de.wikipedia.org/wiki/Monte-Carlo-Simulation> Zugriff 16.07.2015
- <https://de.wikipedia.org/wiki/Varianzreduktion> Zugriff 18.07.2015
- http://www.sitmo.com/article/antithetic_engine_adaptor/ Zugriff 16.07.2015
- https://de.wikipedia.org/wiki/Geschichtete_Zufallsstichprobe Zugriff 23.07.2015
- http://www.fernuni-hagen.de/ksw/neuestatistik/content/files/modul_28433.pdf Zugriff 26.07.2015
- <http://www.glondish.com/blog/> Zugriff 27.07.2015
- http://www.statistik.lmu.de/institut/lehrstuhl/wisoz/lehre/Stichproben_ws1415/download/Stichproben_ws1415_2.pdf Zugriff 27.07.2015

Abbildungsverzeichnis

1	Approximation von Pi	5
2	Vergleich der regulären und antithetischen Methode (bei obigem Beispiel)	13
3	Schichten und Stichproben	14
4	Latin Hypercube Sample im 2-dimensionalen.	17