TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DIPLOMARBEIT

Pricing Asian options by importance sampling

ausgeführt am Institut für
Wirtschaftsmathematik
der Technischen Universität Wien

unter Anleitung von

Dr. Stefan Gerhold

durch

Daniel Ferstl
Milutgasse 2
7400 Oberwart

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Wien, im Mai 2012 ____________________________

Danksagung

Ich möchte mich an dieser Stelle bei all jenen Personen bedanken, die mir bei der Erstellung der Diplomarbeit zur Seite standen. Vor allem gilt mein Dank Herrn Dr. Stefan Gerhold, der mich als Betreuer mit seinen Anregungen und seinem Rat bei angenehmer Atmosphäre hilfreich unterstützte.

Ein großes Dankeschön möchte ich meiner Familie aussprechen, die mich immer unterstützt hat. Ganz besonderer Dank gilt dabei meinen Eltern, Karl und Regina, die mir das Studium ermöglicht haben. Meiner Freundin Laura danke ich dafür, dass sie mich angetrieben hat die Diplomarbeit zügig abzuschließen.

Außerdem möchte ich mich bei meinen Freunden bedanken, die mein Studium zu einem unvergesslichen Abschnitt meines Lebens werden ließen.

Vielen Dank, Daniel

Contents

1 Introduction **6**

2 Theory **7**
2.1 Black-Scholes model 7
2.1.1 Brownian motion 7
2.1.2 Itô's formula 8
2.1.3 Market 9
2.1.4 Assets 9
2.1.5 Change of measure 10
2.1.6 Black-Scholes formula 11
2.2 Asian Options 12
2.2.1 Arithmetic Asian Option 12
2.2.2 Geometric Asian Option 13
2.3 Introduction to Large Deviations 13
2.3.1 Cramér's Theorem for the empirical average 14
2.3.2 The large deviation principle 15
2.3.3 Schilder's Theorem 16
2.4 Variance reduction techniques 18
2.4.1 Importance Sampling 18
2.4.2 The method of control variates 19
2.5 Euler-Lagrange equation 20
2.5.1 The one dimensional Euler-Lagrange equation 21
2.5.2 The Euler-Lagrange equation for a functional with two occurrences of integrals 23
2.6 Price of geometric average Asian options 24

3 Optimal Importance Sampling **28**
3.1 The optimal change of drift 28
3.2 Proof of Theorem 3.1.5 31
3.3 Optimal change of drift for Asian options 36
3.3.1 Optimal change of drift for the geometric average Asian call option 36
3.3.2 Optimal change of drift for the geometric average Asian put option 39
3.3.3 Optimal change of drift for the arithmetic average Asian call option 41
3.3.4 Optimal change of drift for the arithmetic average Asian put option 45

4 Different Monte Carlo estimators **47**
4.1 Monte Carlo estimator without importance sampling 47
4.2 Monte Carlo estimator with importance sampling 48
4.2.1 Using the asymptotically optimal drift of a geometric average Asian option 48
4.2.2 Using the asymptotically optimal drift of an arithmetic average Asian option 49
4.3 Monte Carlo estimator using the method of control variates 50

5 Results **52**
5.1 Arithmetic average Asian call option 52
5.2 Arithmetic average Asian put option 56

Appendix **63**
Maple codes 63

Chapter 1

Introduction

Consider an arithmetic average Asian option. This is a kind of option, whose pay-off depends not just on the value of the underlying at maturity but on all values during the contract period(path-dependent option). Monte Carlo simulation is the method of choice for pricing complex derivatives, such as path-dependent options. The main reason for the popularity of this method is ease of implementation, which only requires the ability to generate sample paths of the asset price and to evaluate the corresponding derivative payoffs.
Now consider the option to be way out-of-money, which means that it is very unlikely for the option to be valuable at maturity. Then an event with small probability accounts for most of the option price. In this case an asymptotic confidence interval given by the central limit theorem can be very unreliable, since even a relatively large sample size can miss rare but large payoffs, generating a low estimate for the payoff combined with a low variance. This means that it is likely to underestimate the value of such an option. Therefore we try to improve the Monte Carlo estimator by using variance reduction techniques such as importance sampling or the method of control variates.
While for using the method of control variates nothing very special has to be wondered about, to use importance sampling one has to derive the change of drift, that minimizes variance. To find the optimal change of drift for Asian options we will use some large deviations techniques, as in [4].
After that we will be able to accomplish the aim of this thesis, which is to compare the results of pricing way out-of-money arithmetic average Asian call respectively put options by using different Monte Carlo estimators. While the case of the call option has already been treated in [4], the case of the put option will be investigated for the first time.
In chapter 2 basic theory is given, such as the Black-Scholes model, Itô's formula, some results of large deviations techniques, variance reduction techniques and the closed form solution for the price of a geometric average Asian option in the Black-Scholes model.
Chapter 3 shows how to derive the optimal change of drift in theory and for the case of geometric and arithmetic average Asian options.
The different Monte Carlo estimators, which will be used to price the option are stated in chapter 4, while the final results are given in the last chapter.
In the appendix one can find the used Maple codes.

Chapter 2

Theory

This chapter provides the basic theory for what will be needed later on. At first we will take a look at the Black-Scholes model, which we will use for determining the price of Asian options. Therefore the second section will be an introduction to this kind of options. The following section gives an overview of large deviations techniques. Combined with importance sampling, what will be investigated in the fourth section, large deviations will help us to reduce variance significantly while trying to price way out-of-money Asian options. In section five some theory about Euler-Lagrange equations is provided, while in the last section we show how to derive the expected payoff of a geometric average Asian option in the Black-Scholes model.

2.1 Black-Scholes model[1]

This section provides a short introduction into the Black-Scholes or Samuelson model as far as it is important for our work later on. We start with the definition of a Brownian motion and present Itô's formula. After that the actual model is presented.

2.1.1 Brownian motion

At first consider some basic definitions to achieve the probability space on which we will define Brownian motions.

Definition 2.1.1. *Let Ω be a non-empty set and $\mathcal{P}(\Omega)$ its power set. A subset $\mathcal{F} \subset \mathcal{P}(\Omega)$ is called σ-algebra with respect to Ω, if it satisfies the following properties:*

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

Definition 2.1.2. *A sequence of σ-algebras $\mathbb{F} = \{\mathcal{F}\}_{t\geq 0}$ is called filtration, if $\forall s,t \geq 0$, $s < t$, it holds that $\mathcal{F}_s \subseteq \mathcal{F}_t$.*

[1] cf. [7] and [9]

A filtration is often used to represent the increasing amount of information one gains by time.

Definition 2.1.3. *A stochastic process $\{X_t\}_{t\geq 0}$ is said to be adapted to the filtration $\{\mathcal{F}\}_{t\geq 0}$ if any random variable X_t is $\mathcal{F}_t$-measurable.*

Definition 2.1.4. *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space. An $\mathbb{R}^d$-valued stochastic process $\{W_t\}_{t\geq 0}$, adapted to $\mathbb{F}$, is called a d-dimensional Brownian motion with respect to $\mathbb{F}$ and $\mathbb{P}$, if it satisfies*

1. *$W_t - W_s$ is independent of $\mathcal{F}_s$, $\forall s, t \in [0, \infty), s < t$ (independence of increments),*
2. *$\forall s, t \in [0, \infty), s < t$, it holds that $(W_{s+t} - W_s) \stackrel{d}{=} (W_t - W_0)$ (stationarity of increments),*
3. *$\forall s, t \in [0, \infty), s < t$, it holds that $W_t - W_s \sim N(0, (t-s)I_d)$,*
4. *$\{W_t\}_{t\geq 0}$ has continuous paths a.s.,*

where I_d is the $(d \times d)$-identity matrix and $\stackrel{d}{=}$ means to have the same distribution.
If additionally $\mathbb{P}(W_0 = 0) = 1$ holds, then $\{W_t\}_{t\geq 0}$ is called a standard Brownian motion.

Note that $\mathbb{P}$ is called *Wiener measure*, the probability law on the space of continuous functions, vanishing at zero and that if $\mathcal{F}_t$ contains the information of $\{W_s\}_{s\in[0,T]}$, $\{\mathcal{F}\}_{t\geq 0}$ is called the *natural filtration* of the Brownian motion.

2.1.2 Itô's formula[2]

This section will recall Itô's formula for the one-dimensional case. The detailed theory about how to derive that result is not provided. For more information about this topic one can have a look at [8].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\mathbb{F} = \{\mathcal{F}\}_{t\geq 0}$, $\{W_t\}_{t\geq 0}$ a one-dimensional $(\mathbb{F}, \mathbb{P})$-Brownian motion and $I \subset [0, \infty)$ an interval of the form $I = [a, b]$, $I = [a, b)$ or $I = [a, \infty)$ with $a < b$.

Definition 2.1.5. *Let $\mathbb{W}_{\mathbb{F}}(I)$ be the set of all functions $f : I \times \Omega \to \mathbb{R}$ satisfying*

1. *f is progressively measurable, i.e. $f|_{[a,t]\times\Omega}$ is $\mathcal{B}([a, t]) \otimes \mathcal{F}_t$-measurable $\forall\ t \in I$,*
2. *$\mathbb{P}(\int_a^t f^2(s, \omega)ds < \infty\ \forall\ t \in I) = 1$.*

Now we can define a special kind of process.

Definition 2.1.6. *Let $v \in \mathbb{W}_{\mathbb{F}}([0, \infty))$ and $u : [0, \infty) \times \Omega \to \mathbb{R}$ be progressively measurable satisfying $\mathbb{P}(\int_0^t |u(s, \cdot)|\, ds < \infty,\ \forall\ t \geq 0) = 1$. Let X_0 be $\mathcal{F}_0$-measurable. Then*

$$X_t(\omega) := X_0(\omega) + \int_0^t u(s, \omega)ds + \int_0^t v(s, \omega)dW_s(\omega),\ t \geq 0,\ \omega \in \Omega$$

is called Itô-process. An abbreviated version is given by

$$dX_t = u(t)dt + v(t)dW_t.$$

[2] cf. [8]

Now we can formulate:

Theorem 2.1.7 (Itô-formula)**.** *Let $U \subset \mathbb{R}$ be open and $\{X_t\}_{t\geq 0}$ an Itô-process with values in U. Let $g : [0,\infty) \times U \to \mathbb{R}$ be continuous differentiable once with respect to the first argument and twice with respect to the second argument($g \in C^{1,2}$) and let these partial derivatives be continuous on $[0,\infty) \times U$.*
Then $Y_t := g(t, X_t),\ t \geq 0$, is an Itô-process and for almost all $\omega \in \Omega$ it holds that

$$\begin{aligned} Y_t(\omega) &= g(0, X_0(\omega)) \\ &\quad + \int_0^t \left(\frac{\partial g}{\partial s}(s, X_s(\omega)) + \frac{\partial g}{\partial x}(s, X_s(\omega))\, u(s,\omega) + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(s, X_s(\omega))\, v^2(s,\omega) \right) ds \\ &\quad + \int_0^t \frac{\partial g}{\partial x}(s, X_s(\omega))\, v(s,\omega)\, dW_s(\omega),\ t \geq 0. \end{aligned} \tag{2.1}$$

An abbreviated version is given by

$$dY_t = \frac{\partial g}{\partial t}(t, X_t(\omega))\, dt + \frac{\partial g}{\partial x}(t, X_t(\omega))\, dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t(\omega))\, (dX_t)^2,$$

with $(dt)^2 = 0,\ dt dW_t = 0 = dW_t dt$ and $(dW_t)^2 = dt$.

Now we can begin to present the actual Black-Scholes model.

2.1.3 Market

At first we have to mention, that we have to make several assumptions on the financial market, to be able to derive the Black-Scholes model:

- Trading is possible in continuous time.
- There are no trading restrictions.
- Interest rates for lending and borrowing money are equal.
- There are no costs or taxes.

This is called a complete market.

2.1.4 Assets

In this model there exist two types of assets, a riskless and a risky one. The price of the *riskless asset(bond)* is denoted by B_t and is described by the ordinary differential equation

$$dB_t = rB_t dt, \tag{2.2}$$

where the constant r determines the instantaneous interest rate.
Let $B_0 = 1$, then $B_t = B_0 e^{rt}$ for $t \geq 0$ solves 2.2.

Let$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, where Ω is a non-empty set, $\mathcal{F}$ a σ-algebra, $\mathbb{F} = \{\mathcal{F}\}_{t\geq 0}$ a filtration of $\mathcal{F}$ and $\mathbb{P}$ a probability measure.

Then the price of the *risky asset(stock)* S_t is described by the stochastic differential equation

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \tag{2.3}$$

where $\mu \in \mathbb{R}$ is the appreciation rate, $\sigma > 0$ the volatility and $\{W_t\}_{t>0}$ a one-dimensional standard $(\mathbb{F}, \mathbb{P})$-Brownian motion. The part in front of dt (in this case μ) is called *drift* of the risky asset.

Using Itô's formula one can proof that $S_t = S_0 e^{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t}$ is a solution to 2.3. S_t is called a geometric Brownian motion.

$\tilde{S}_t = \frac{S_t}{B_t} = e^{-rt} S_t = S_0 e^{\sigma W_t + (\mu - r - \frac{1}{2}\sigma^2)t}$ determines the discounted price of the risky asset and is described by the stochastic differential equation

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = (\mu - r)dt + \sigma dW_t. \tag{2.4}$$

2.1.5 Change of measure

The intention of the model is now to find a probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, so that $\{\tilde{S}_t\}_{t \in [0,T]}$, $T > 0$, becomes a martingale with respect to the new measure. Thereby equivalent means that $\mathbb{Q}$ has the same null sets as $\mathbb{P}$.

At first consider some definitions and theorems:

Definition 2.1.8. *Let T be an index set. A stochastic process $\{M_t\}_{t \in [0,T]}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called martingale with respect to the filtration $\{\mathcal{F}_t\}_{t \in [0,T]}$, if*

1. *$\{M_t\}_{t \in [0,T]}$ is adapted to $\{\mathcal{F}_t\}_{t \in [0,T]}$,*
2. *$\mathbb{E}[|M_t|] < \infty$, $\forall t \in [0, T]$,*
3. *$\mathbb{E}[M_t | \mathcal{F}_s] = M_s$, $\forall s, t \in [0, T]$, $s < t$ (martingale property).*

The special case of the Radon-Nikodym theorem for probability measures states the following

Theorem 2.1.9. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathbb{Q}$ a probability measure equivalent to $\mathbb{P}$. Then there exists an integrable random variable Z with $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$ and $\mathbb{E}[Z] = 1$.*

Last but not least we formulate Girsanov's theorem, which will be very important for the rest of the section.

Theorem 2.1.10 (Girsanov's theorem)**.** *Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space with $\mathbb{F} = \{\mathcal{F}_t\}_{0 \le t \le T}$ the natural filtration of the standard $(\mathbb{F}, \mathbb{P})$-Brownian motion $\{W_t\}_{0 \le t \le T}$. Let $(\theta_t)_{0 \le t \le T}$ be an adapted process satisfying $\int_0^T \theta_s^2 ds < \infty$ and that the process $\{L_t\}_{0 \le t \le T}$, defined by*

$$L_t = \exp\left(- \int_0^T \theta_s dW_s - \frac{1}{2} \int_0^T \theta_s^2 ds \right),$$

is a martingale. Consider the probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, with Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(- \int_0^T \theta_s dW_s - \frac{1}{2} \int_0^T \theta_s^2 ds \right),$$

then $\{W_t^*\}_{0 \le t \le T}$ with $W_t^* = W_t + \int_0^t \theta_s ds$ is a $(\mathbb{F}, \mathbb{Q})$-standard Brownian motion.

Now recall that we want to find a probability measure $\mathbb{Q}$ that makes $\{\tilde{S}_t\}_{t\in[0,T]}$ a martingale. In the Black-Scholes model there exists an explicit measure satisfying this condition. A Radon-Nikodym derivative is given by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^T \frac{r-\mu}{\sigma}dW_t - \frac{1}{2}\int_0^T \frac{(r-\mu)^2}{\sigma^2}dt\right) = \exp\left(\frac{r-\mu}{\sigma}W_T - \frac{1}{2}\frac{(r-\mu)^2}{\sigma^2}T\right). \tag{2.5}$$

Because of Girsanov's Theorem we know, that we receive a new standard $(\mathbb{F},\mathbb{Q})$-Brownian motion $\{W_t^*\}_{t\in[0,T]}$, where $W_t^* = W_t - \int_0^t \frac{r-\mu}{\sigma} = W_t - \frac{r-\mu}{\sigma}t, t \in [0,T]$. With respect to $\mathbb{Q}$, the discounted price of the risky asset $\{\tilde{S}_t\}_{t\in[0,T]}$ is a martingale and is described by the stochastic differential equation

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = \sigma dW_t^*. \tag{2.6}$$

Using Itô's formula one can proof that $\tilde{S}_t = S_0 e^{\sigma W_t^* - \frac{1}{2}\sigma^2 t}$ is a solution to 2.6.

The ordinary price of the risky asset S_t with respect to $\mathbb{Q}$ is described by the stochastic differential equation

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t^*. \tag{2.7}$$

The solution to 2.7 is given by $S_t = S_0 e^{\sigma W_t^* + (r-\frac{1}{2}\sigma^2)t}$.

As one can see in 2.7 the drift of the risky asset changed from μ to r by changing the probability measure. Later on we will use a variance reduction technique called importance sampling, which takes advantage of the fact that changing the probability measure also changes the drift.
Note that $\mathbb{Q}$ is called the risk-neutral measure. If we talk about S_t under the risk-neutral measure, S_t is always meant to be like the solution to 2.7. In this case the appreciation rate μ has no impact on the price of the risky asset anymore, S_t only depends on the volatility σ and the instantaneous interest rate r.

2.1.6 Black-Scholes formula

This subsection provides the Black-Scholes formula for pricing an European option, since we will need it to derive a closed form solution for the price of the geometric average Asian option in the Black-Scholes model.

Theorem 2.1.11 (Black-Scholes formula)**.** *Consider a European call option with strike $K \geq 0$ and maturity $T > 0$. Then the price of the option at time $t \in [0,T]$ is given by*

$$C_t = C_t(S_t, T-t, K, r, \sigma) = S_t\Phi(d_1(S_t, T-t, K, r, \sigma)) - Ke^{-r(T-t)}\Phi(d_2(S_t, T-t, K, r, \sigma)), \tag{2.8}$$

where Φ denotes the distribution function of the standard normal distribution and

$$d_{1,2}(S_t, T-t, K, r, \sigma) = \frac{\log(\frac{S_t}{K}) + (r \pm \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}.$$

Consider now a European put option. Then the price is given by

$$P_t = P_t(S_t, T-t, K, r, \sigma) = Ke^{-r(T-t)}\Phi(-d_2(S_t, T-t, K, r, \sigma)) - S_t\Phi(-d_1(S_t, T-t, K, r, \sigma)). \tag{2.9}$$

2.2 Asian Options[3]

Asian options belong to the group of so-called "exotic options", meaning that they do not have that big impact on the market, though Asian options are the most popular options among this group. The difference between Asian options and their European counterparts is that the pay-off does not just depend on the value of the underlying asset at the maturity date, but on an average of all values during the contract period. Therefore one has to simulate the whole path and not just a single value, if one wants to estimate the price of an Asian option by doing a Monte Carlo simulation.

2.2.1 Arithmetic Asian Option

The average value of the underlying asset in discrete respectively continuous time is determined by the arithmetic average

$$\overline{S} = \frac{1}{n}\sum_{i=1}^{n} S_{t_i},$$
$$\overline{S} = \frac{1}{T}\int_0^T S_t,$$

where T is the duration of the contract, $t_1 < t_2 < ... < t_n = T$ are the associated trading dates in discrete time and S_t the price of the underlying of the option at time t.

Therefore the pay-off of an arithmetic Asian call respectively put option is given by

$$\left(\overline{S} - K\right)^+,$$
$$\left(K - \overline{S}\right)^+,$$

where K determines the strike price.

Thus the price of the option at time 0 is denoted by

$$e^{-rT}\mathbb{E}\left[\left(\overline{S} - K\right)^+\right],$$
$$e^{-rT}\mathbb{E}\left[\left(K - \overline{S}\right)^+\right].$$

[3] cf. [10]

2.2.2 Geometric Asian Option

The average value of the underlying asset in discrete time is determined by the geometric average

$$\hat{S} = \left(\prod_{i=1}^{n} S_{t_i} \right)^{\frac{1}{n}}, \tag{2.10}$$

where $t_1 < t_2 < ... < t_n = T$ are the associated trading dates in discrete time, T the duration of the contract and S_t the price of the underlying of the option at time t.

Since 2.10 can be rewritten as

$$\hat{S} = e^{\ln\left(\left(\prod_{i=1}^{n} S_{t_i}\right)^{\frac{1}{n}} \right)} = e^{\frac{1}{n}\sum_{i=1}^{n} ln\left(S_{t_i}\right)},$$

the average value of the underlying asset in continuous time is determined by

$$\hat{S} = e^{\frac{1}{T}\int_0^T \ln(S_t)}.$$

Therefore the pay-off of an arithmetic Asian call respectively put option is

$$(\hat{S} - K)^+,$$
$$(K - \hat{S})^+,$$

where K determines the strike price.
Thus the price of the option at time 0 is denoted by

$$e^{-rT}\mathbb{E}\left[\left(\hat{S} - K \right)^+ \right],$$
$$e^{-rT}\mathbb{E}\left[\left(K - \hat{S} \right)^+ \right].$$

2.3 Introduction to Large Deviations[4]

Large deviations theory is a part of probability theory that deals with the description of so-called *rare events*, where rare means that random variables differ from its mean by more than a "normal" amount. Normal usually means what is described by the central limit theorem.
The area of large deviations covers a set of asymptotic results on rare event probabilities and a set of methods to derive such results.
Among other topics large deviations find important applications in finance, where rare events play an important role. Approximations, done for pricing options, in particular for pricing barrier options and way out-of-money options, are good examples and also in the scope of this thesis. Therefore large deviations techniques aim to quantify the probability of rare events on exponential scale.

[4] cf. [1] and [2]

Let us start with a short example:
Consider a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$ and let $X_1, X_2, ...$ be i.i.d. random variables with

$$\mathbb{E}X_1 = \mu \in \mathbb{R},$$
$$VarX_1 = \sigma^2 \in (0, \infty)$$

and let $S_n = X_1 + \cdots + X_n$ $(n \in \mathbb{N})$ be the partial sums.

There are two fundamental theorems dealing with such sequences:

Strong Law of Large Numbers(SLLT)

$$\frac{1}{n} S_n \xrightarrow{n \to \infty} \mu \quad \mathbb{P}\text{-a.s.}$$

Central Limit Theorem(CLT)

$$\frac{1}{\sigma\sqrt{n}}(S_n - \mu n) \xrightarrow{n \to \infty} Z \quad \text{in law w.r.t. } \mathbb{P},$$

where Z is a standard normal random variable.
The CLT quantifies the probability that S_n differs from μn by an amount of order $\sqrt{n}$, what we call a "normal" deviation.
Events where S_n differs from μn by an amount of order n lead to deviations that are called "large" (large deviations).
As an example consider the following event

$$\{S_n \geq (\mu + a)n\}, \ a > 0,$$

whose probability tends to zero as $n \to \infty$. The question now is to determine how quick this happens. Therefore our task is to quantify the rate at which the probability tends to zero.

2.3.1 Cramér's Theorem for the empirical average

Theorem 2.3.1. [5] *Let (X_i) be i.i.d. $\mathbb{R}$-valued random variables with $\phi(t) = \mathbb{E}e^{tX_1} < \infty$, $\forall t \in \mathbb{R}$, where ϕ denotes the generating function.*
Let $S_n = \sum_{i=1}^n X_i$. Then for all $a > \mathbb{E}X_1$,

$$\lim_{n \to \infty} \frac{1}{n} \ln(\mathbb{P}(S_n \geq an)) = -I(a), \tag{2.11}$$

where

$$I(z) = \sup_{t \in \mathbb{R}} (zt - \ln(\phi(t))), \ z \in \mathbb{R} \tag{2.12}$$

is called a rate function.

[5] cf. [1, Theorem I.4]

Note that 2.11 also holds for $\mathbb{P}(S_n \leq an)$ with $a < \mathbb{E}X_1$, that $I(z)$ is called the Fenchel-Legendre transform of $\ln \phi(t)$ and that $\ln \phi$ is the cumulant generating function.
Assuming that the conditions of Cramér's Theorem are satisfied, the rate function 2.12 has the following properties[6]:

1. I is lower semi-continuous and convex on $\mathbb{R}$.
2. I has compact level sets.
3. I is continuous and strictly convex on $\text{int}(\mathcal{D}_\mathcal{I})$, where $\mathcal{D}_\mathcal{I} = \{z \in \mathbb{R} : \ I(z) < \infty\}$ and $\text{int}(\mathcal{D}_\mathcal{I})$ is the interior of $\mathcal{D}_\mathcal{I}$.
4. I is smooth on $\text{int}(\mathcal{D}_\mathcal{I})$.
5. $I(z) \geq 0$ with equality if and only if $z = \mu$.
6. $I''(\mu) = \frac{1}{\sigma^2}$

Remarks:

- The level sets of I are the sets $I^{-1}([0, c]) = \{z \in \mathbb{R} : I(z) \leq c\}$ with $c \in [0, \infty)$.
- Lower semi-continuity is equivalent to the level sets being closed.
- The convexity of I implies, that $\mathcal{D}_\mathcal{I}$ is an interval (possibly infinite).

2.3.2 The large deviation principle

Now we do not consider i.i.d. random variables any longer, but formulate a more general theory.

Let $\mathcal{X}$ be a Polish space, i.e. a separable completely metrizable topological space, with distance(metric) $d : \ \mathcal{X} \times \mathcal{X} \to [0, \infty)$.

Definition 2.3.2. $f : \mathcal{X} \to [-\infty, \infty]$ *is lower semi-continuous if it satisfies any of the following equivalent properties:*

1. $\liminf\limits_{n\to\infty} f(x_n) \geq f(x)$ *for all* (x_n), x *such that* $x_n \to x$ *in* $\mathcal{X}$.

2. $\lim\limits_{\epsilon \downarrow 0} \inf\limits_{y \in B_\epsilon(x)} f(y) = f(x)$ *with* $B_\epsilon(x) = \{y \in \mathcal{X} : \ d(x, y) < \epsilon\}$.

3. f *has closed level sets, i.e.* $f^{-1}([-\infty, c]) = \{x \in \mathcal{X} : \ f(x) \leq c\}$ *is closed for all* $c \in \mathbb{R}$.

Now we can define a rate function.

Definition 2.3.3. $I : \mathcal{X} \to [0, \infty]$ *is a rate function if:*

1. $I \not\equiv \infty$

[6] cf. [1, exercise I.16]

2. *I is lower semi-continuous.*

3. *I has compact level sets.*

Theorem 2.3.4. [7] *A lower semi-continuous function achieves a minimum on every non-empty compact set.*

Definition 2.3.5. *A sequence of probability measures $(\mathbb{P}_n)$ on $\mathcal{X}$ is said to satisfy the* ***large deviation principle****(LDP) with rate n and rate function I if*

1. *I is a rate function in the sense of Definition 2.3.3.*

2. $\limsup\limits_{n\to\infty} \frac{1}{n} \ln \mathbb{P}_n(C) \leq -I(C) \quad \forall C \subset \mathcal{X}$ *closed.*

3. $\liminf\limits_{n\to\infty} \frac{1}{n} \ln \mathbb{P}_n(O) \geq -I(O) \quad \forall O \subset \mathcal{X}$ *open.*

where for $S \subseteq \mathcal{X}$ $I(S)$ is defined as $I(S) = \inf\limits_{x\in S} I(x)$.

Theorem 2.3.6. [8] *Let $(\mathbb{P}_n)$ satisfy the LDP. Then the associated rate function I is unique.*

The next result, known as Varadhan's Lemma, is the extension of the Laplace approximation for integrals to a general (infinite-dimensional) setting.

Lemma 2.3.7 (Varadhan's Lemma)**.** [9] *Let $(\mathbb{P}_n)$ satisfy the LDP on $\mathcal{X}$ with rate n and rate function I. Let $H : \mathcal{X} \to \mathbb{R}$ be continuous and bounded from above.Then*

$$\lim_{n\to\infty} \frac{1}{n} \ln \int_{\mathcal{X}} e^{nH(x)} d\mathbb{P}_n(x) = \sup_{x\in\mathcal{X}} [H(x) - I(x)] \tag{2.13}$$

2.3.3 Schilder's Theorem

In this section we will take a closer look to a more specific result, namely *Schilder's Theorem* for Sample Path Large Deviations.

Let $(W_t)_{t\in[0,T]}$ denote a standard Brownian motion in $\mathbb{R}^d$. Consider the process

$$W_\epsilon(t) = \sqrt{\epsilon} W_t.$$

Theorem 2.3.8. [10] *For any integer d and any $\tau, \epsilon, \delta > 0$,*

$$\mathbb{P}\left(\sup_{0\leq t\leq \tau} \|W_\epsilon(t)\| \geq \delta \right) \leq 4d \exp\left(\frac{-\delta^2}{2d\tau\epsilon} \right). \tag{2.14}$$

[7] cf. [1, exercise III.4]
[8] cf. [1]
[9] cf. [1] or [2, Theorem 4.3.1]
[10] cf. [2, Lemma 5.2.1]

Now let
$H_0([0,T]) = \{f : [0,T] \to \mathbb{R} : f$ is absolutely continuous on $[0,T],\ f' \in L_2,\ f(0) = 0\}$ denote the space of all absolutely continuous functions with square integrable derivative, vanishing at zero.
Let μ_ϵ be the probability measure induced by $W_\epsilon(\cdot)$ on $C[0,T]$, the space of all continuous functions, vanishing at zero, equipped with the supremum norm topology.

Theorem 2.3.9 (Schilder)**.** [11] *$\{\mu_\epsilon\}$ satisfies an LDP on $C[0,T]$ with rate function*

$$I(h) = \begin{cases} \frac{1}{2}\int_0^T \|\dot{h}(t)\|^2 dt, & \text{if } h \in H_0 \\ \infty & \text{otherwise} \end{cases} \tag{2.15}$$

Let us show the lower bound of this LDP.

Proof. Consider G a nonempty open set of $C([0,T])$, $h \in G$ and $\delta > 0$ s.t. $B(h,\delta) \subset G$. We want to prove that

$$\liminf_{\epsilon \to 0} \epsilon \ln \mathbb{P}\left[\sqrt{\epsilon} W \in B(h,\delta)\right] \geq -I(h).$$

For $h \notin H_0([0,T])$, this inequality is trivial since $I(h) = \infty$. Suppose now $h \in H_0([0,T])$, and consider the probability measure:

$$\frac{d\mathbb{Q}^h}{d\mathbb{P}} = \exp\left(\int_0^T \frac{\dot{h}_t}{\sqrt{\epsilon}} dW_t - \frac{1}{2\epsilon}\int_0^T \left|\dot{h}_t\right|^2 dt\right),$$

so that by Girsanov's Theorem, $W^h := W - \frac{h}{\sqrt{\epsilon}}$ is a Brownian motion under $\mathbb{Q}^h$. Then, we have

$$\begin{aligned}
\mathbb{P}\left[\sqrt{\epsilon} W \in B(h,\delta)\right] &= \mathbb{P}\left[\left|W^h\right| < \frac{\delta}{\sqrt{\epsilon}}\right] \\
&= \mathbb{E}_{\mathbb{Q}^h}\left[\exp\left(-\int_0^T \frac{\dot{h}_t}{\sqrt{\epsilon}} dW_t^h - \frac{1}{2\epsilon}\int_0^T \left|\dot{h}_t\right|^2 dt\right) \mathbb{1}_{|W^h| < \frac{\delta}{\sqrt{\epsilon}}}\right] \\
\left(W^h \mathbb{Q}^h - BM\right) &= \mathbb{E}\left[\exp\left(-\int_0^T \frac{\dot{h}_t}{\sqrt{\epsilon}} dW_t - \frac{1}{2\epsilon}\int_0^T \left|\dot{h}_t\right|^2 dt\right) \mathbb{1}_{|W| < \frac{\delta}{\sqrt{\epsilon}}}\right] \\
\left(W \tilde{-} W\right) &= \mathbb{E}\left[\exp\left(+\int_0^T \frac{\dot{h}_t}{\sqrt{\epsilon}} dW_t - \frac{1}{2\epsilon}\int_0^T \left|\dot{h}_t\right|^2 dt\right) \mathbb{1}_{|W| < \frac{\delta}{\sqrt{\epsilon}}}\right] \\
&= \mathbb{E}\left[\exp\left(-\frac{1}{2\epsilon}\int_0^T \left|\dot{h}_t\right|^2 dt\right) \cosh\left(\int_0^T \frac{\dot{h}_t}{\sqrt{\epsilon}} dW_t\right) \mathbb{1}_{|W| < \frac{\delta}{\sqrt{\epsilon}}}\right] \\
&\geq \exp\left(-\frac{1}{2\epsilon}\int_0^T \left|\dot{h}_t\right|^2 dt\right) \mathbb{P}\left[|W| < \frac{\delta}{\sqrt{\epsilon}}\right].
\end{aligned}$$

This implies

$$\epsilon \ln \mathbb{P}\left[\sqrt{\epsilon} W \in B(h,\delta)\right] \geq -I(h) + \epsilon \ln \mathbb{P}\left[|W| < \frac{\delta}{\sqrt{\epsilon}}\right],$$

and thus the required lower bound. □

[11] cf. [2, Theorem 5.2.3]

Since it is much more complicated to show the upper bound, the proof is not given here. For more information one can have a look at [2, Lemma 5.2.3].

2.4 Variance reduction techniques

There are several different types of so-called variance reduction techniques, but all of them intend to increase the efficiency of Monte Carlo methods by reducing the variance of simulation estimates. In the following sections we will take a closer look at two of them. At first the method called *importance sampling* is presented, followed by the method of *control variates.*

2.4.1 Importance sampling[12]

Importance sampling attempts to reduce variance by changing the probability measure from which paths are generated. The idea is to try to give more weight to "important" outcomes of the simulation. For example if one investigates barrier options, one can make it more probable that the price of an asset exceeds or falls below a specific value.
Thereby the expected value will be unchanged under the new measure, but the variance will "hopefully" decrease.

Now let us have a look at a short example:
Let

$$\mu = \mathbb{E}[h(X)] = \int h(x)f(x)dx \tag{2.16}$$

be the term we want to estimate, where X is a random variable of $\mathbb{R}^d$ with probability density f and h is a function from $\mathbb{R}^d$ to $\mathbb{R}$. Therefore we get the following Monte Carlo estimator for n independent draws $X_1, ..., X_n$ of f:

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

Let g be any other probability density on $\mathbb{R}^d$ satisfying $f(x) > 0 \Rightarrow g(x) > 0,\ \forall x \in \mathbb{R}^d$. Then we can rewrite 2.16 as

$$\mu = \int h(x)\frac{f(x)}{g(x)}g(x)dx = \tilde{\mathbb{E}}\left[h(X)\frac{f(X)}{g(X)}\right], \tag{2.17}$$

where $\tilde{\mathbb{E}}$ indicates that the expectation is taken with X distributed according to g and $\frac{f(x)}{g(x)}$ is a *Radon-Nikodym derivative.* Therefore we get a new Monte Carlo estimator, this time for n independent draws $X_1, ..., X_n$ of g:

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

With 2.17 we see that $\tilde{\mathbb{E}}[\hat{\mu}_g] = \mu$ and thus that $\hat{\mu}_g$ is an unbiased estimator of μ.
Since the expected value is equal with and without Importance Sampling, the difference

[12] cf. [3, section 4.6]

of the variance has to be found by comparing the second moments. With importance sampling, we get

$$\tilde{\mathbb{E}}\left[\left(h(X)\frac{f(X)}{g(X)}\right)^2\right] = \mathbb{E}\left[h(X)^2\frac{f(X)}{g(X)}\right].$$

Since the second moment without importance sampling is $\mathbb{E}[h(X)^2]$, one can see, that the change of variance depends on the choice of g. Therefore achieving a rise of variance is as possible as achieving a reduction. So we see that the success of this method lies in the art of selecting an effective density g.

Note the following:
Consider the special case in which h is nonnegative. Then $h(x)f(x)$ is also nonnegative and may be normalized to a probability density. Then

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

would be the perfect choice, leading to a zero-variance estimator $\hat{\mu}_g$. Unfortunately $\int h(x)f(x)dx = \mu$ is what we wanted to estimate in the first place. Therefore in practice we try to find an approximately optimal g.

This was an easy example to illustrate the idea of importance sampling. Later on, when we calculate the price of an Asian option in the Black-Scholes model, we will take advantage of this method. Since underlying assets then are represented by geometric Brownian motions, changing the probability measure will result in changing the drift of the Brownian motion.

2.4.2 The method of control variates[13]

Now we take a look at another variance reduction technique, the method of control variates. It exploits information about the errors in estimates of known quantities to reduce the error in an estimate of an unknown quantity.
For example let $Y_1, \ldots, Y_n$ be the payoffs of an option with respect to sample path i. Supposing that all Y_i are i.i.d., the usual estimator of $\mathbb{E}[Y_i]$ is the sample mean $\bar{Y} = (Y_1 + \cdots + Y_n)/n$. This estimator is unbiased and converges with probability 1 as $n \to \infty$.
Suppose now that for each sample path we compute the output X_i along with Y_i. Suppose that the pairs (X_i, Y_i) are i.i.d. for $i = 1, \ldots, n$ and the expectation $\mathbb{E}[X]$ of the X_i is known.(We use (X,Y) to denote a pair of random variables with the same distribution as each (X_i, Y_i).) Then for any fixed b we can calculate

$$Y_i(b) = Y_i - b(X_i - \mathbb{E}[X])$$

from the ith sample path and then compute the sample mean

$$\bar{Y}(b) = \bar{Y} - b(\bar{X} - \mathbb{E}[X]) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - b(X_i - \mathbb{E}[X])). \tag{2.18}$$

[13] cf. [3, section 4.1]

This is a control variate estimator; the observed error $\bar{X} - \mathbb{E}[X]$ serves as a control in estimating $\mathbb{E}[Y]$. Since

$$\mathbb{E}[\bar{Y}(b)] = \mathbb{E}[\bar{Y} - b(\bar{X} - \mathbb{E}[X])] = \mathbb{E}[\bar{Y}] = \mathbb{E}[Y]$$

the control variate estimator 2.18 is unbiased and it is consistent, with probability 1, because of

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Y_i(b) &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (Y_i - b(X_i - \mathbb{E}[X])) \\
&= \mathbb{E}[Y - b(X - \mathbb{E}[X])] \\
&= \mathbb{E}[Y]
\end{aligned}$$

Each $Y_i(b)$ has the following variance

$$\begin{aligned}
Var(Y_i(b)) &= Var(Y_i - b(X_i - \mathbb{E}[X])) && (2.19) \\
&= \sigma_Y^2 - 2b\sigma_X\sigma_Y\rho_{XY} + b^2\sigma_X^2 \equiv \sigma^2(b), && (2.20)
\end{aligned}$$

where ρ_{XY} is the correlation between X and Y, $\sigma_X^2 = Var(X)$ and $\sigma_Y^2 = Var(Y)$. The control variate estimator $\bar{Y}(b)$ has variance σ_Y^2/n and the ordinary sample mean $\bar{Y}(b = 0)$ has variance σ_Y^2/n. Thus the control variate estimator has smaller variance than the standard one if $b^2\sigma_X^2 < 2b\sigma_X\sigma_Y\rho_{XY}$.
The optimal coefficient b^* minimizes the variance 2.20 and is given by

$$b^* = \frac{\sigma_Y}{\sigma_X}\rho_{XY} = \frac{Cov(X,Y)}{Var(X)}.$$

Since in practice it is unlikely that σ_Y or ρ_{XY} is known, if $\mathbb{E}[Y]$ is unknown, we have to estimate b^*. Therefore we get

$$\hat{b}_n = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

Now we got everything to use this method.

Since in the Black-Scholes model there exists a closed form solution for the price of an geometric average Asian option, we can use this price as a control variate for the price of an arithmetic average Asian option.

2.5 Euler-Lagrange equation[14]

Since later on we will have to determine the *Euler-Lagrange equation*(or *Euler's equation*), this section provides what we need to know.
The Euler-Lagrange equation belongs to the field of *Variations of Functionals*, where the concept of the variation(or differential) of a functional is analogous to the concept of a

[14] cf. [6, Chapter 1]

differential of a function.
Consider a functional of the form

$$J[y] = \int_a^b L(x, y(x), y'(x))dx, \tag{2.21}$$

and let

$$\Delta J[h] = J[y+h] - J[y]$$

be its increment, corresponding to the increment $h = h(x)$ of the "independent variable" $y = y(x)$. If y is fixed, $\Delta J[h]$ is a functional of h. Suppose that

$$\Delta J[h] = \phi[h] + \epsilon\|h\|, \tag{2.22}$$

where $\phi[h]$ is a linear functional and $\epsilon \to 0$ as $\|h\| \to 0$. Then the functional $J[y]$ is said to be *differentiable* and the *principal linear part* of the increment $\Delta J[h]$, i.e. the linear functional $\phi[h]$ which differs from $\Delta J[h]$ by an infinitesimal of order higher than 1 relative to $\|h\|$, is called the *variation*(or differential) of $J[h]$ and is denoted by $\delta J[h]$.

Problems involving the determination of maxima and minima of functionals are called *variational problems*. Consider the "simplest" variational problem, which is formulated as follows:

. *Let $L(x, y, y')$ be a function with continuous first and second (partial) derivatives with respect to all its arguments. Then, among all the functions y(x) which are continuously differentiable for $a \leq x \leq b$ and satisfy the boundary conditions*

$$y(a) = A, \qquad y(b) = B,$$

find the function for which the functional

$$J[y] = \int_a^b L(x, y(x), y'(x))dx \tag{2.23}$$

has a weak extremum.

In other words, the simplest variational problem consists of finding the weak extremum of a functional of the form 2.23. This can be done by solving Euler's equation 2.24, which is denoted as follows.

2.5.1 The one dimensional Euler-Lagrange equation

Theorem 2.5.1 (Euler's equation)**.** [15] *Let $J[y]$ be a functional of the form*

$$J[y] = \int_a^b L(x, y(x), y'(x))dx,$$

defined on the set of functions $y(x)$ which have continuous first derivatives in $[a, b]$ and satisfy the boundary conditions $y(a) = A,\ y(b) = B$. Then a necessary condition for $J[y]$ to have an extremum for a given function $y(x)$ is that $y(x)$ satisfy Euler's equation

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'} = 0. \tag{2.24}$$

[15] cf. [6, Theorem 1]

Proof. Suppose we give $y(x)$ an increment $h(x)$ with $y_\epsilon(x) = y(x) + \epsilon h(x)$ and $y'_\epsilon(x) = y'(x) + \epsilon h'(x)$, where $h(x)$ is a differentiable function and ϵ is small. Since we want the function y_ϵ to continue to satisfy the boundary conditions, we must have h to satisfy $h(a) = h(b) = 0$.
Define

$$J_\epsilon[y] = \int_a^b L(x, y_\epsilon, y'_\epsilon)dx = \int_a^b L_\epsilon dx,$$

where $L_\epsilon = L(x, y_\epsilon, y'_\epsilon)$.
The total derivative of a function f of several variables x, y, z with respect to one of its input variables, e.g. x, is given by

$$\frac{df}{dx} = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx}.$$

Now we calculate the total derivative of $J_\epsilon[y]$ with respect to ϵ, which is

$$\frac{dJ_\epsilon}{d\epsilon} = \frac{d}{d\epsilon}\int_a^b L_\epsilon dx = \int_a^b \frac{dL_\epsilon}{d\epsilon}dx.$$

With

$$\frac{dL_\epsilon}{d\epsilon} = \frac{\partial L_\epsilon}{\partial \epsilon} + \frac{\partial L_\epsilon}{\partial x}\frac{dx}{d\epsilon} + \frac{\partial L_\epsilon}{\partial y_\epsilon}\frac{dy_\epsilon}{d\epsilon} + \frac{\partial L_\epsilon}{\partial y'_\epsilon}\frac{dy'_\epsilon}{d\epsilon},$$

where

$$\frac{\partial L_\epsilon}{\partial \epsilon} = \frac{\partial}{\partial \epsilon}L(x, y_\epsilon, y'_\epsilon) = 0, \text{ since } L \text{ does not depend explicitly on } \epsilon,$$
$$\frac{\partial L_\epsilon}{\partial x}\frac{dx}{d\epsilon} = 0, \text{ since } \frac{dx}{d\epsilon} = 0,$$
$$\frac{dy_\epsilon}{d\epsilon} = h \text{ and}$$
$$\frac{dy'_\epsilon}{d\epsilon} = h'$$

and with $\epsilon = 0$ we get

$$\frac{dJ_\epsilon}{d\epsilon} = \int_a^b \left[h(x)\frac{\partial L_\epsilon}{\partial y_\epsilon} + h'(x)\frac{\partial L_\epsilon}{\partial y'_\epsilon}\right]dx = \int_a^b \left[h(x)\frac{\partial L}{\partial y} + h'(x)\frac{\partial L}{\partial y'}\right]dx. \tag{2.25}$$

According to [6, Sec. 3.2, Theorem 2] a necessary condition for $J[y]$ to have an extremum for $y = y(x)$ is that

$$\frac{dJ_\epsilon}{d\epsilon} = 0,$$

therefore we get

$$\int_a^b \left[h(x)\frac{\partial L}{\partial y} + h'(x)\frac{\partial L}{\partial y'}\right]dx = 0. \tag{2.26}$$

Applying the following lemma to 2.26 yields the Euler-Lagrange equation.

Lemma 2.5.2. [16] *If $\alpha(x)$ and $\beta(x)$ are continuous in $[a,b]$ and if*

$$\int_a^b [\alpha(x)h(x) + \beta(x)h'(x)]dx = 0$$

for every function $h(x)$, defined on $[a,b]$, continuous and with continuous first derivatives, such that $h(a) = h(b) = 0$, then $\beta(x)$ is differentiable and $\beta'(x) = \alpha(x)$, $\forall x \in [a,b]$.

□

2.5.2 The Euler-Lagrange equation for a functional with two occurrences of integrals

This was the standard case. Now the method of the proof will be used once again to derive the Euler-Lagrange equation for a functional with two occurrences of integrals. Therefore consider the functions $L_1(x,y,y')$ and $L_2(x,y,y')$ as L before and let $J_1[y] = \int_a^b L_1(x,y,y')dx$ and $J_2[y] = \int_a^b L_2(x,y,y')dx$ be functionals of the form as $J[y]$. Define y_ϵ, y'_ϵ, h and h' as before and let $J_{1,\epsilon}[y] = \int_a^b L_{1,\epsilon}dx$ and $J_{2,\epsilon}[y] = \int_a^b L_{2,\epsilon}dx$, where $L_{i,\epsilon} = L_i(x, y_\epsilon, y'_\epsilon)$, $i = 1,2$.

Now define $F[J_1, J_2]$ as a functional of the two integrals J_1 and J_2. For this functional we want to find the corresponding Euler-Lagrange equation. Therefore we calculate the total derivative of $F_\epsilon = F[J_{1,\epsilon}, J_{2,\epsilon}]$ with respect to ϵ, which is given by

$$\frac{dF_\epsilon}{d\epsilon} = \frac{\partial F_\epsilon}{\partial \epsilon} + \frac{\partial F_\epsilon}{\partial J_{1,\epsilon}}\frac{dJ_{1,\epsilon}}{d\epsilon} + \frac{\partial F_\epsilon}{\partial J_{2,\epsilon}}\frac{dJ_{2,\epsilon}}{d\epsilon}, \tag{2.27}$$

where

$$\frac{\partial F_\epsilon}{\partial \epsilon} = 0, \text{ since } F_\epsilon \text{ does not depend explicitly on } \epsilon.$$

Since $J_{1,\epsilon}$ and $J_{2,\epsilon}$ are of the form of J_ϵ in the standard case, we get, as in 2.25,

$$\begin{aligned} \frac{dJ_{1,\epsilon}}{d\epsilon} &= \int_a^b \left[h(x)\frac{\partial L_{1,\epsilon}}{\partial y_\epsilon} + h'(x)\frac{\partial L_{1,\epsilon}}{\partial y'_\epsilon}\right] dx, \\ \frac{dJ_{2,\epsilon}}{d\epsilon} &= \int_a^b \left[h(x)\frac{\partial L_{2,\epsilon}}{\partial y_\epsilon} + h'(x)\frac{\partial L_{2,\epsilon}}{\partial y'_\epsilon}\right] dx. \end{aligned}$$

Using integration by parts for the second term of the integral yields

$$\int_a^b h'(x)\frac{\partial L}{\partial y'}dx = h(x)\frac{\partial L}{\partial y'}\bigg|_a^b - \int_a^b h(x)\frac{d}{dt}\frac{\partial L}{\partial y'}dx = -\int_a^b h(x)\frac{d}{dt}\frac{\partial L}{\partial y'}dx,$$

where the last equality holds because $h(a) = h(b) = 0$. Applying this result gives

$$\begin{aligned} \frac{dJ_{1,\epsilon}}{d\epsilon} &= \int_a^b h(x)\left(\frac{\partial L_{1,\epsilon}}{\partial y_\epsilon} - \frac{d}{dt}\frac{\partial L_{1,\epsilon}}{\partial y'_\epsilon}\right) dx, \\ \frac{dJ_{2,\epsilon}}{d\epsilon} &= \int_a^b h(x)\left(\frac{\partial L_{2,\epsilon}}{\partial y_\epsilon} - \frac{d}{dt}\frac{\partial L_{2,\epsilon}}{\partial y'_\epsilon}\right) dx. \end{aligned}$$

[16] cf. [6, Lemma 4]

Substituting this into 2.27 yields

$$\begin{aligned}\frac{dF_\epsilon}{d\epsilon} &= \frac{\partial F_\epsilon}{\partial J_{1,\epsilon}}\int_a^b h(x)\left(\frac{\partial L_{1,\epsilon}}{\partial y_\epsilon}-\frac{d}{dt}\frac{\partial L_{1,\epsilon}}{\partial y'_\epsilon}\right)dx+\frac{\partial F_\epsilon}{\partial J_{2,\epsilon}}\int_a^b h(x)\left(\frac{\partial L_{2,\epsilon}}{\partial y_\epsilon}-\frac{d}{dt}\frac{\partial L_{2,\epsilon}}{\partial y'_\epsilon}\right)dx\\ &= \int_a^b h(x)\left(\frac{\partial F_\epsilon}{\partial J_{1,\epsilon}}\left(\frac{\partial L_{1,\epsilon}}{\partial y_\epsilon}-\frac{d}{dt}\frac{\partial L_{1,\epsilon}}{\partial y'_\epsilon}\right)+\frac{\partial F_\epsilon}{\partial J_{2,\epsilon}}\left(\frac{\partial L_{2,\epsilon}}{\partial y_\epsilon}-\frac{d}{dt}\frac{\partial L_{2,\epsilon}}{\partial y'_\epsilon}\right)\right)dx\end{aligned}$$

According to [6, Sec. 3.2, Theorem 2] a necessary condition for $F[J_1, J_2]$ to have an extremum for $y = y(x)$ is that

$$\frac{dF_\epsilon}{d\epsilon} = 0,$$

hence we get for $\epsilon = 0$

$$\frac{dF_\epsilon}{d\epsilon} = \int_a^b h(x)\left(\frac{\partial F}{\partial J_1}\left(\frac{\partial L_1}{\partial y}-\frac{d}{dt}\frac{\partial L_1}{\partial y'}\right)+\frac{\partial F}{\partial J_2}\left(\frac{\partial L_2}{\partial y}-\frac{d}{dt}\frac{\partial L_2}{\partial y'}\right)\right)dx = 0.$$

Since this integral has to be equal to zero for any increment h, it is the rest of the integrand who has to be equal to zero. Therefore we get

$$\frac{\partial F}{\partial J_1}\left(\frac{\partial L_1}{\partial y}-\frac{d}{dt}\frac{\partial L_1}{\partial y'}\right)+\frac{\partial F}{\partial J_2}\left(\frac{\partial L_2}{\partial y}-\frac{d}{dt}\frac{\partial L_2}{\partial y'}\right) = 0,$$

what is equvialent to the following Euler-Lagrange equation

$$\frac{\frac{\partial L_2}{\partial y}-\frac{d}{dx}\frac{\partial L_2}{\partial y'}}{\frac{\partial L_1}{\partial y}-\frac{d}{dx}\frac{\partial L_1}{\partial y'}} = -\frac{\frac{\partial F}{\partial J_1}}{\frac{\partial F}{\partial J_2}}. \tag{2.28}$$

2.6 Price of geometric average Asian options[17]

As mentioned before there exists a closed form solution for the price of geometric average Asian options in the Black-Scholes model. Since we will need this expectation later on to use the method of control variates, the formula for the price is provided in this section.

At first note that for arithmetic average Asian options there is no such solution. The reason why the geometric case is easier to handle, is that the product, other than the sum, of log-normal distributed random variables is also log-normal distributed. This of course is very helpful, since, as we know, the geometric Brownian motion S_t is log-normal distributed.
The payoff function for the discrete geometric average Asian call option is given by

$$\left(\exp\left(\frac{1}{m+1}\sum_{i=0}^{m}\log(S(t_i))\right)-K\right)^+ = \left(\left(\prod_{i=0}^{m}S(t_i)\right)^{1/(m+1)}-K\right)^+,$$

[17] cf. [11, Section 3.2]

where m is the number of time points and $0 = t_0 < \ldots < t_m = T$ with $t_i = \frac{iT}{m}$. Note that

$$\begin{aligned}\prod_{i=0}^{m} S(t_i) = & \frac{S(t_m)}{S(t_{m-1})}(\frac{S(t_{m-1})}{S(t_{m-2})})^2(\frac{S(t_{m-2})}{S(t_{m-3})})^3 \\ & \cdots(\frac{S(t_3)}{S(t_2)})^{m-2}(\frac{S(t_2)}{S(t_1)})^{m-1}(\frac{S(t_1)}{S(t_0)})^m S_0^{m+1}.\end{aligned}$$

Since $S_t = S_0 e^{\sigma W_t^* + (r - \frac{1}{2}\sigma^2)t}$ and $t_i - t_{i-1} = \frac{T}{m},\ i = 1, \ldots, m$ it follows that

$$\begin{aligned}\frac{S(t_m)}{S(t_{m-1})} &= \exp(\sigma(W_{t_m}^* - W_{t_{m-1}}^*) + \left(r - \frac{\sigma^2}{2}\right)(t_m - t_{m-1}) \\ &= \exp(\sigma\sqrt{T/m}X_1 + \left(r - \frac{\sigma^2}{2}\right)T/m), \\ \frac{S(t_{m-1})}{S(t_{m-2})} &= \exp(\sigma\sqrt{T/m}X_2 + \left(r - \frac{\sigma^2}{2}\right)T/m), \\ &\vdots \\ \frac{S(t_1)}{S(t_0)} &= \exp(\sigma\sqrt{T/m}X_m + \left(r - \frac{\sigma^2}{2}\right)T/m),\end{aligned}$$

where $\{X_i\}_{i=1}^m$ are independent, standard normal distributed random variables.

With the above results we get

$$
\begin{aligned}
\log\left(\frac{\left(\prod_{i=0}^m S\left(t_i\right)\right)^{1/(m+1)}}{S_0}\right) &= \log\left(\left(\frac{\prod_{i=0}^m S\left(t_i\right)}{S_0^{m+1}}\right)^{1/(m+1)}\right) \\
&= \frac{1}{m+1}\log\left(\frac{\prod_{i=0}^m S\left(t_i\right)}{S_0^{m+1}}\right) \\
&= \frac{1}{m+1}\log\left(\frac{S\left(t_m\right)}{S\left(t_{m-1}\right)}\left(\frac{S\left(t_{m-1}\right)}{S\left(t_{m-2}\right)}\right)^2\cdots\left(\frac{S\left(t_2\right)}{S\left(t_1\right)}\right)^{m-1}\left(\frac{S\left(t_1\right)}{S\left(t_0\right)}\right)^m\right) \\
&= \frac{1}{m+1}\Bigg(\log\left(\frac{S\left(t_m\right)}{S\left(t_{m-1}\right)}\right)+2\log\left(\frac{S\left(t_{m-1}\right)}{S\left(t_{m-2}\right)}\right)+ \\
&\cdots+(m-1)\log\left(\frac{S\left(t_2\right)}{S\left(t_1\right)}\right)+m\log\left(\frac{S\left(t_1\right)}{S\left(t_0\right)}\right)\Bigg) \\
&= \frac{1}{m+1}\Bigg(\left(\sigma\sqrt{T/m}X_1+\left(r-\frac{\sigma^2}{2}\right)T/m\right) \\
&+2\left(\sigma\sqrt{T/m}X_2+\left(r-\frac{\sigma^2}{2}\right)T/m\right)+ \\
&\cdots+(m-1)\left(\sigma\sqrt{T/m}X_{m-1}+\left(r-\frac{\sigma^2}{2}\right)T/m\right) \\
&+m\left(\sigma\sqrt{T/m}X_m+\left(r-\frac{\sigma^2}{2}\right)T/m\right)\Bigg) \\
&= \frac{1}{m+1}\left(\sigma\sqrt{T/m}\sum_{i=1}^m iX_i+(1+2+\cdots+m)\left(r-\frac{\sigma^2}{2}\right)\frac{T}{m}\right) \\
&= \frac{\sigma\sqrt{T/m}\sum_{i=1}^m iX_i}{m+1}+\frac{\left(r-\frac{\sigma^2}{2}\right)T}{2}.
\end{aligned}
\tag{2.29}
$$

Caused by the additive mean and variance of independent normal random variables, we know that

$$
\frac{\sigma\sqrt{T/m}\sum_{i=1}^m iX_i}{m+1}\sim N\left(0,\underbrace{\frac{\sigma^2T(1^2+2^2+\ldots+m^2)}{m(m+1)^2}}_{\frac{(2m+1)\sigma^2T}{6(m+1)}}\right).
$$

Therefore we have

$$
\frac{\sigma\sqrt{T/m}\sum_{i=1}^m iX_i}{m+1}=\sigma\sqrt{\frac{(2m+1)T}{6(m+1)}}Z, \tag{2.30}
$$

where Z is standard normal distributed.

Plugging 2.30 in 2.29 we get

$$\log\left(\frac{\left(\prod_{i=0}^{m} S(t_i)\right)^{1/(m+1)}}{S_0}\right) = \left(\rho - \frac{\sigma_Z^2}{2}\right) T + \sigma_Z \sqrt{T} Z,$$

where $\sigma_Z = \sigma\sqrt{\frac{2m+1}{6(m+1)}}$ and $\rho = \frac{\left(r-\frac{\sigma^2}{2}\right)+\sigma_Z^2}{2}$.
Hence, we can obtain the price of the geometric average Asian call option:

$$\begin{aligned}
C_0^g\left(S_0, T, K, r, \sigma\right) &= \exp\left(-rT\right)\mathbb{E}\left[\left(\left(\prod_{i=0}^{m} S\left(t_i\right)\right)^{1/(m+1)} - K\right)^+\right] \\
&= \exp\left(\left(\rho - r\right)T\right)\exp\left(-\rho T\right)\mathbb{E}\left[\left(S_0\exp\left(\left(\rho - \frac{\sigma_Z^2}{2}\right)T + \sigma_Z\sqrt{T}Z\right) - K\right)^+\right] \\
&= \exp\left(\left(\rho - r\right)T\right)C_0\left(S_0, T, K, \rho, \sigma_Z\right),
\end{aligned}$$

where $C_0(S_0, T, K, \rho, \sigma_Z)$ is the price of an European call option with interest rate ρ and volatility σ_Z.

By the Black-Scholes formula 2.8 we get

$$C_0(S_0, T, K, \rho, \sigma_Z) = S_0\Phi(d_1(S_0, T, K, \rho, \sigma_Z)) - Ke^{-\rho T}\Phi(d_2(S_0, T, K, \rho, \sigma_Z)),$$

where Φ denotes the distribution function of the normal distribution and

$$d_{1,2}\left(S_0, T, K, \rho, \sigma_Z\right) = \frac{\log\left(\frac{S_0}{K}\right) + \left(\rho \pm \frac{\sigma_Z^2}{2}\right)T}{\sqrt{T}\sigma_Z}.$$

Therefore we have

$$\begin{aligned}
C_0^g(S_0, T, K, r, \sigma) &= \exp((\rho - r)T)(S_0\Phi(d_1(S_0, T, K, \rho, \sigma_Z)) - Ke^{-\rho T}\Phi(d_2(S_0, T, K, \rho, \sigma_Z))) \\
&= \exp(-rT)(S_0e^{\rho T}\Phi(d_1(S_0, T, K, \rho, \sigma_Z)) - K\Phi(d_2(S_0, T, K, \rho, \sigma_Z)))
\end{aligned} \tag{2.31}$$

Let P respectively P^g be the price of the corresponding European respectively geometric average Asian put option. Then the price of the European one is given by

$$P_0(S_0, T, K, \rho, \sigma_Z) = Ke^{-\rho T}\Phi(-d_2(S_0, T, K, \rho, \sigma_Z)) - S_0\Phi(-d_1(S_0, T, K, \rho, \sigma_Z)).$$

Hence we have

$$\begin{aligned}
P_0^g(S_0, T, K, r, \sigma) &= \exp((\rho - r)T)(Ke^{-\rho T}\Phi(-d_2(S_0, T, K, \rho, \sigma_Z)) - S_0\Phi(-d_1(S_0, T, K, \rho, \sigma_Z))) \\
&= \exp(-rT)(K\Phi(-d_2(S_0, T, K, \rho, \sigma_Z)) - S_0e^{\rho T}\Phi(-d_1(S_0, T, K, \rho, \sigma_Z))).
\end{aligned} \tag{2.32}$$

Chapter 3

Optimal Importance Sampling[1]

This chapter provides the theory Paolo Guasoni and Scott Robertson used in their paper to derive the optimal change of drift of an underlying asset in the Black-Scholes model, where optimal is meant in the sense of reducing variance via importance sampling, as well as the practical example for a geometric respectively an arithmetic average Asian option.

3.1 The optimal change of drift[2]

Consider the Black-Scholes model and let $\mathbb{P}$ determine the risk-neutral probability measure. Then the risky asset is, as mentioned before,

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

where W_t is a standard Brownian motion under $\mathbb{P}$, r the interest rate and σ the volatility.

Now we want to describe the payoff of a derivative by a functional depending on the shocks process $(W_t)_{t\in[0,T]}$. Therefore denote by

$$\mathbb{W}_T \equiv \{x \in C([0,T], \mathbb{R}) : x(0) = 0\}$$

the Wiener space of continuous functions on $[0,T]$ vanishing at zero.
This space is endowed with the topology of uniform convergence and with the usual Wiener measure $\mathbb{P}$, defined on the completion of the Borel σ-field $\mathcal{F}_T$, under which the coordinate process $W_t(x) = x_t$ is a standard Brownian motion with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, the usual augmentation of the natural filtration of W.
Roughly spoken, $\mathbb{W}_T$ contains the paths of $(W_t)_{t\in[0,T]}$.

Definition 3.1.1. *A payoff is a non-negative functional $G : \mathbb{W}_T \to \mathbb{R}_+$, continuous in the uniform topology.*

[1] cf. [4]
[2] cf. [4, chapter 3]

Example 3.1.2. *Consider the arithmetic average Asian call option. Its payoff is given by* $\left(\frac{1}{T}\int_0^T S_t dt - K\right)^+$*, which corresponds to the functional*

$$G(x) = \left(\frac{1}{T}\int_0^T S_0 e^{\sigma x_t + (r - \frac{1}{2}\sigma^2)t} dt - K\right)^+ .$$

Let $F = \log G$, taking values in $\mathbb{R} \cup \{-\infty\}$, and define by

$$\mathbb{H}_T \equiv \{h \in AC[0,T] : h(0) = 0, \int_0^T \dot{h}_t^2 dt < \infty\}$$

the Cameron-Martin space of absolutely continuous functions vanishing at zero with square integrable derivative.

Now for any deterministic $h \in \mathbb{H}_T$, the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}^h}{d\mathbb{P}} = \exp(\int_0^T \dot{h}_t dW_t - \frac{1}{2}\int_0^T \dot{h}_t^2 dt)$$

induces an equivalent probability measure $\mathbb{Q}^h$. Under $\mathbb{Q}^h$ the process $W_t^* \equiv W_t - h_t$ is a standard Brownian motion as we learned in section 2.1.5.
So we see that $\mathbb{H}_T$ contains the possible changes of drift of S_t.

Since we want to minimize

$$Var_{\mathbb{Q}^h}\left(G\frac{d\mathbb{P}}{d\mathbb{Q}^h}\right) = \mathbb{E}_{\mathbb{Q}^h}\left[\left(G\frac{d\mathbb{P}}{d\mathbb{Q}^h}\right)^2\right] - \mathbb{E}_{\mathbb{Q}^h}\left[G\frac{d\mathbb{P}}{d\mathbb{Q}^h}\right]^2 = \mathbb{E}_{\mathbb{P}}\left[G^2\frac{d\mathbb{P}}{d\mathbb{Q}^h}\right] - \mathbb{E}_{\mathbb{P}}\left[G\right]^2,$$

we have to minimize

$$\mathbb{E}_{\mathbb{P}}\left[G^2\frac{d\mathbb{P}}{d\mathbb{Q}^h}\right] = \mathbb{E}_{\mathbb{P}}\left[e^{2F(W) - \int_0^T \dot{h}_t dW_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt}\right].$$

When Monte Carlo simulation is necessary to estimate $\mathbb{E}_{\mathbb{P}}[G]$, the above quantity is, in general, intractable.

Instead, as in [5], we consider the small-noise asymptotics

$$L(h) = \limsup_{\epsilon \downarrow 0} \epsilon \log\left(\mathbb{E}_{\mathbb{P}}\left[e^{\frac{1}{\epsilon}\left(2F(\sqrt{\epsilon}W) - \int_0^T \sqrt{\epsilon}\dot{h}_t dW_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)}\right]\right),$$

which correspond to approximating $Var_{\mathbb{Q}^h}\left(G\frac{d\mathbb{P}}{d\mathbb{Q}^h}\right)$ with $e^{L(h)}$. So instead of minimizing the variance, we minimize $e^{L(h)}$, particularly $L(h)$. Note that since we use this approximation, the minimizer h will "just" be asymptotically optimal.

Definition 3.1.3. *An asymptotically optimal drift is a solution to the problem*

$$\min_{h \in \mathbb{H}_T} L(h).$$

Now we have to find a deterministic expression for $L(h)$, which becomes suitable for optimization. This is possible under the following

Assumption 3.1.4. $F : \mathbb{W}_T \to \mathbb{R} \cup \{-\infty\}$ *is continuous and satisfies*

$$F(x) \leq K_1 + K_2 \max_{t\in[0,T]} |x_t|^\alpha \tag{3.1}$$

for some constants $K_1,\ K_2 > 0$ *and* $\alpha \in (0,2)$.

Note that condition 3.1 is fulfilled for virtually all options of practical interest.

Theorem 3.1.5. *Let F satisfy Assumption 3.1.4. Then:*

1. *If $h \in \mathbb{H}_T$, and $\dot{h}$ has finite variation, then*

$$L(h) = \sup_{x\in\mathbb{H}_T} \left(2F(x) + \frac{1}{2}\int_0^T (\dot{x}_t - \dot{h}_t)^2 dt - \int_0^T \dot{x}_t^2 dt \right) \tag{3.2}$$

2. *For all $h \in \mathbb{H}_T$, there exist maximizers to both 3.2 and 3.3 below:*

$$\sup_{x\in\mathbb{H}_T} \left(2F(x) - \int_0^T \dot{x}_t^2 dt \right) \tag{3.3}$$

3. *If $\hat{h}$ is a solution to 3.3, then $\hat{h}$ is asymptotically optimal if*

$$L(\hat{h}) = 2F(\hat{h}) - \int_0^T \dot{\hat{h}}_t^2 dt. \tag{3.4}$$

   Furthermore, if 3.4 holds, then $\hat{h}$ is the unique solution of 3.3.

A proof is provided in the next section. This Theorem yields that if Assumption 3.1.4 is satisfied, an asymptotically optimal drift $\hat{h}$ can be determined by solving 3.3, particularly by solving the corresponding Euler-Lagrange ordinary differential equation of 3.3. If this $\hat{h}$ has a derivative with finite variation and satisfies condition 3.4, then it actually is an asymptotically optimal drift.
Note that condition 3.4 certainly holds when F is a concave functional, since then we have a unique maximum, but in general, one has to solve a new variational problem to evaluate $L(\hat{h})$, which also reduces to an Euler-Lagrange ODE.

Once $\hat{h}$ is found, and since

$$dW_t = d(W_t^* + \hat{h}_t) = dW_t^* + \dot{\hat{h}}_t dt \tag{3.5}$$

we achieve

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[G] &= \mathbb{E}_{\mathbb{Q}^{\hat{h}}}\left[G \frac{d\mathbb{P}}{d\mathbb{Q}^{\hat{h}}}\right] \\
&= \mathbb{E}_{\mathbb{Q}^{\hat{h}}}\left[e^{F(W)} e^{-\int_0^T \dot{\hat{h}}_t dW_t + \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}\right] \\
&= \mathbb{E}_{\mathbb{Q}^{\hat{h}}}\left[e^{F\left(W^*+\hat{h}\right) - \int_0^T \dot{\hat{h}}_t dW_t^* - \int_0^T \dot{\hat{h}}_t^2 dt + \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}\right] \\
&= \mathbb{E}_{\mathbb{Q}^{\hat{h}}}\left[e^{F\left(W^*+\hat{h}\right) - \int_0^T \dot{\hat{h}}_t dW_t^* - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}\right] \\
&= \mathbb{E}_{\mathbb{Q}^{\hat{h}}}\left[G\left(W^* + \hat{h}\right) e^{-\int_0^T \dot{\hat{h}}_t dW_t^* - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}\right],
\end{aligned} \tag{3.6}$$

where W^* is a standard Brownian motion under $\mathbb{Q}^{\hat{h}}$.

As one can see, in a new Monte Carlo simulation the payoff has to be rescaled by the factor $e^{-\int_0^T \dot{\hat{h}}_t dW_t^* - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}$, while with 3.5 we get

$$\begin{aligned}
&\frac{dS_t}{S_t} = rdt + \sigma dW_t \\
\Leftrightarrow\ &\frac{dS_t}{S_t} = rdt + \sigma(dW_t^* + \dot{\hat{h}}_t dt) \\
\Leftrightarrow\ &\frac{dS_t}{S_t} = (r + \sigma\dot{\hat{h}}_t)dt + \sigma dW_t^*.
\end{aligned}$$

for the stochastic differential equation of the risky asset S_t. Here we can see that the drift of S_t changes from r to $r + \sigma\dot{\hat{h}}_t$. Using Itô's formula one can proof that a solution is given by $S_t = S_0 e^{\sigma(W_t^* + \hat{h}_t) + (r - \frac{1}{2}\sigma^2)t}$.

3.2 Proof of Theorem 3.1.5[3]

The proof of Theorem 3.1.5 is divided into several lemmas.The first one shows the existence of solutions to the problems 3.2 and 3.3, using a standard variational argument.

Lemma 3.2.1. *Let F satisfy Assumption 3.1.4. Then for any $h \in \mathbb{H}_T$ and $M > 0$ there exists a maximizer for the problem*

$$\max_{x\in\mathbb{H}_T}\left(2F(x) + M\int_0^T(\dot{x}_t - \dot{h}_t)^2 dt - 2M\int_0^T \dot{x}_t^2 dt + (1-2M)\int_0^T \dot{h}_t^2 dt\right). \tag{3.7}$$

Proof. Recall that if $\dot{g}_n \to \dot{g}$ weakly in $L^2[0,T]$, then $g_n \to g$ uniformly in $[0,T]$. Since F is continuous in the uniform norm, it follows that it is also weakly continuous. Let $M > 0$ and fix $h \in \mathbb{H}_T$. Rewrite 3.7 as

$$\max_{x\in\mathbb{H}_T}\left(2F(x) - M\|h + x|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2\right). \tag{3.8}$$

As a function of x, $M\|h + x|_{\mathbb{H}_T}^2$ is convex and finite, hence norm-continuous. Thus, it is also weakly lower semi-continuous. Since F is weakly continuous, the function $x \to 2F(x) - M\|h + x|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2$ is then weakly upper semi-continuous.
Assumption 3.1.4 implies that

$$\begin{aligned}
2F(x) - M\|h + x|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2 &\le 2K_1 + 2K_2\|x\|_\infty^\alpha - M\|h + x|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2 \\
&\le 2K_1 + 2K_2 T^{\alpha/2}\|x\|_{\mathbb{H}_T}^\alpha - M\|h + x|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2.
\end{aligned}$$

Since $\alpha < 2$, the coercivity property follows, i.e.

$$\lim_{\|x\|_{\mathbb{H}_T}\to\infty}\left(2F(x) - M\|h + x|_{\mathbb{H}_T}^2 + \|h\|_{\mathbb{H}_T}^2\right) = -\infty,$$

and the existence of a maximizer follows by upper semi-continuity. □

[3] cf. [4, Appendix]

The remaining part of the proof of Theorem 3.1.5 now requires the theory on large deviations, which we treated in section 2.3. Therefore recall Schilder's Theorem 2.15 and Varadhan's Lemma 2.13 and consider them adjusted for our case:

Theorem 3.2.2 (Schilder)**.** *Let $X = \mathbb{W}_T$ and μ_ϵ be the probability on $\mathbb{W}_T$ induced by the process $\sqrt{\epsilon}W$, where W is a standard Brownian motion. Then $\{(\mu_\epsilon\}_{\epsilon\in(0,\delta)}$ satisfies an LDP with rate function*

$$I(x) = \begin{cases} \frac{1}{2}\int_0^T \|\dot{x}(t)\|^2 dt, & if\, x \in \mathbb{H}_T \\ \infty & if x \in \mathbb{W}_T \backslash \mathbb{H}_T \end{cases} \tag{3.9}$$

Lemma 3.2.3 (Varadhan's Lemma)**.** *Let $(Z_\epsilon)_{\epsilon\in(0,\delta)}$ be a family of X-valued random variables, whose laws μ_ϵ satisfy a large deviations principle rate function I. If $H : \mathcal{X} \to \mathbb{R}$ is a continuous function which satisfies*

$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[exp\left(\frac{\gamma}{\epsilon}H(Z_\epsilon)\right)\right] < \infty, \tag{3.10}$$

for some $\gamma > 1$, then

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right)\right] = \sup_{x\in X}[H(x) - I(x)] \tag{3.11}$$

The following Lemma states a slight generalization of the result of Varadhan's Lemma in order to allow $H : X \to [-\infty, \infty)$ instead of $H : X \to \mathbb{R}$.

Lemma 3.2.4. [4] *Let $H : \mathcal{X} \to [-\infty, \infty)$. Under the assumptions of Varadhan's Lemma 2.13, the following holds for any $A \in \mathcal{B}$:*

$$\begin{aligned} \sup_{x\in A^\circ} (H(x) - I(x)) &\leq \liminf_{\epsilon\to 0} \epsilon \log \left(\int_{A^\circ} \exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right) d\mu_\epsilon\right) \\ &\leq \limsup_{\epsilon\to 0} \epsilon \log \left(\int_{\bar{A}} \exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right) d\mu_\epsilon\right) \\ &\leq \sup_{x\in\bar{A}} (H(x) - I(x)). \end{aligned}$$

Proof. The second inequality is trivial, while the first one is the result of [2, Lemma 4.3.4], fixing $x \in A^\circ$ instead of $x \in \mathcal{X}$.
For the third inequality, note that if $H \equiv -\infty$ the result holds trivially. Assuming the other case, let C be a closed subset of $\mathcal{X}$. For $M > 0$, consider the set $C_M = C \cap \{H(x) \geq -M\}$, which is closed by the continuity of H. Thus, one has that

$$\begin{aligned} \int_C exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right) d\mu_\epsilon &= \int_{C_M} exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right) d\mu_\epsilon + \int_{C\backslash C_M} exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right) d\mu_\epsilon \\ &\leq \int_{C_M} exp\left(\frac{1}{\epsilon}H(Z_\epsilon)\right) d\mu_\epsilon + exp\left(-\frac{M}{\epsilon}\right) \mu_\epsilon(C\backslash C_M) \end{aligned}$$

[4] cf. [5, Lemma 2.1]

Since $(\mu_\epsilon)_{\epsilon\in(0,\delta)}$ satisfy the LDP with rate function I,

$$\limsup_{\epsilon\to 0} \epsilon \log\left(exp\left(-\frac{M}{\epsilon}\right)\mu_\epsilon(C\backslash C_M)\right) \le -M - \inf_{x\in\overline{C\backslash C_M}} I(x).$$

Using Varadhan's Lemma 2.13 on $H\mathbb{1}_{C_M}$ as in [2, Exercise 4.3.11],

$$\limsup_{\epsilon\to 0} \epsilon \log\left(\int_{C_M} exp\left(\frac{1}{\epsilon}H(Z_\epsilon)d\mu_\epsilon\right)\right) \le \sup_{x\in C_M}\left(H(x)-I(x)\right) \tag{3.12}$$

and hence [2, Lemma 1.2.15]

$$\begin{aligned}
&\limsup_{\epsilon\to 0} \epsilon \log\left(\int_{C_M} exp\left(\frac{1}{\epsilon}H\left(Z_\epsilon\right)\right)d\mu_\epsilon + exp\left(-\frac{M}{\epsilon}\right)\mu_\epsilon\left(C\backslash C_M\right)\right)\\
&\le \max\left(\sup_{x\in C_M}\left(H\left(x\right)-I\left(x\right)\right), -M - \inf_{x\in\overline{C\backslash C_M}} I\left(x\right)\right)\\
&\le \max\left(\sup_{x\in C}\left(H\left(x\right)-I\left(x\right)\right), -M\right).
\end{aligned}$$

The claim follows, as $M\to\infty$. $\square$

Lemma 3.2.5. *Let F satisfy Assumption 3.1.4, and define $F_h : \mathbb{W}\to\mathbb{R}$ as*

$$F_h(x) = 2F(x) - \int_0^T \dot{h}_t dx_t + \frac{1}{2}\int_0^T \dot{h}_t^2.$$

Then F_h is well-defined, norm-continuous and satisfies 3.10 for any $h\in\mathbb{H}_T$ and $\gamma>1$.

Proof. Since F is continuous, the continuity of F_h will follow from the continuity of $x\mapsto\int_0^T \dot{h}dx_t$. Since $\dot{h}$ has finite variation on $[0,T]$ for each $h\in\mathbb{H}_T$, the integral $\int_0^T \dot{h}dx_t$ is defined path-wise in the Stieltjes sense. For any $f\in\mathbb{W}_T$, integration by parts and the continuity of f imply that

$$\left|\int_0^T \dot{h}df_t\right| = \left|\dot{h}(T)f(T) - \int_0^T f_t d\dot{h}_t\right| \le \|f\|_{\mathbb{W}_T} Var(\dot{h}),$$

where $Var(\dot{h})$ denotes the total variation of $\dot{h}$. Thus, continuity follows by the finite variation assumption. To check the integrability condition 3.10, apply the Cauchy-Schwarz inequality to see that

$$\begin{aligned}
&\epsilon\log\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{\gamma}{\epsilon}\left(2F\left(\sqrt{\epsilon}W\right) - \int_0^T \dot{h}_t d\left(\sqrt{\epsilon}W\right)_t + \frac{1}{2}\int_0^T \dot{h}_t^2 dt\right)\right)\right]\\
&\le \frac{\gamma}{2}\int_0^T \dot{h}_t^2 dt + \frac{\epsilon}{2}\log\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{2\gamma}{\sqrt{\epsilon}}\int_0^T \dot{h}_t dW_t\right)\right]\\
&\quad + \frac{\epsilon}{2}\log\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{4\gamma}{\epsilon}F\left(\sqrt{\epsilon}W\right)\right)\right].
\end{aligned} \tag{3.13}$$

The first term is finite. For the second, observe that $\int_0^T \dot{h}_t dW_t \sim N\left(0, \int_0^T \dot{h}_t^2 dt\right)$, whence

$$\limsup_{\epsilon \to 0} \frac{\epsilon}{2} \log \mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{2\gamma}{\sqrt{\epsilon}} \int_0^T \dot{h}_t dW_t\right)\right] = \gamma^2 \int_0^T \dot{h}_t^2 dt < \infty$$

It remains to consider the last term in 3.13. Assumption 3.4 implies that

$$\frac{\epsilon}{2}\log\left(\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{4\gamma}{\epsilon}F(\sqrt{\epsilon}x)\right)\right]\right) \leq 2\gamma K_1 + \frac{\epsilon}{2}\log\left(\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\left(\sup_{0\leq t\leq T}|W(t)|\right)^{\alpha}\right)\right]\right),$$

and one has to check that the last term is finite. To see this, observe that

$$\begin{aligned} E_{\mathbb{P}}\left[\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\left(\sup_{0\leq t\leq T}|W(t)|\right)^{\alpha}\right)\right] &\leq 2E_{\mathbb{P}}\left[\exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}\left(\sup_{0\leq t\leq T} x(t)\right)^{\alpha}\right)\right] \\ &\leq 4\sqrt{\frac{2}{\pi T}}\int_0^\infty \exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}b^\alpha - \frac{1}{2T}b^2\right)db, \end{aligned}$$

where the first inequality follows from the formula $\mathbb{E}_{\mathbb{P}}[X] = \int_0^\infty \mathbb{P}(X \geq b)db$, combined with the elementary estimate

$$\mathbb{P}(\sup_{0\leq t\leq T}|W(t)| \geq b) \leq 2\mathbb{P}(\sup_{0\leq t\leq T} W(t) \geq b).$$

The second inequality follows from the classical distribution

$$\mathbb{P}(\sup_{0\leq t\leq T} W(t) \in db) = \sqrt{\frac{2}{\pi T}}\exp\left(-\frac{b^2}{2T}\right)db.$$

Applying Lemma 3.2.6 below, for $A = \frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}$, $B = \frac{1}{2T}$, yields

$$\begin{aligned} &\int_0^\infty \exp\left(\frac{4\gamma}{\epsilon^{1-\alpha/2}}K_2 b^\alpha - \frac{1}{2T}b^2\right)db \\ &\leq \exp\left(\frac{4\gamma K_2}{\epsilon^{1-\alpha/2}}(4\gamma K_2\alpha T)^{\frac{\alpha}{2-\alpha}}\epsilon^{-\alpha/2} + \frac{1}{2T}(4\gamma K_2\alpha T)^{\frac{2}{2-\alpha}}\epsilon^{-1}\right) \\ &\times\left((4\gamma K_2\alpha T)^{\frac{1}{2-\alpha}}\epsilon^{-1/2} + \sqrt{\frac{2\pi}{\min\left(\frac{1}{T}, \frac{1}{T}(2-\alpha)\right)}}\right), \end{aligned}$$

which, after letting $N = 4\gamma K_2\alpha T$ and $M = \min(\frac{1}{T}, \frac{1}{T}(2-\alpha))$, reduces to

$$\exp\left(\frac{1}{\epsilon}\frac{1}{T}N^{\frac{2}{2-\alpha}}\left(\frac{1}{\alpha} - \frac{1}{2}\right)\right)\left(N^{\frac{1}{2-\alpha}}\epsilon^{-1/2} + \sqrt{\frac{2\pi}{M}}\right).$$

Thus,

$$\begin{aligned} &\limsup_{\epsilon\to 0}\epsilon\log\left(2\sqrt{\frac{2}{\pi T}}\int_0^\infty \exp\left(\frac{4\gamma}{\epsilon^{1-\alpha/2}}K_2 b^\alpha - \frac{1}{2T}b^2\right)db\right)^{1/2} \\ &\leq \frac{1}{2}\frac{1}{T}N^{\frac{2}{2-\alpha}}\left(\frac{1}{\alpha} - \frac{1}{2}\right) < \infty, \end{aligned}$$

which proves the claim. □

Lemma 3.2.6. *Let A, $B > 0$, $\alpha \in (0,2)$ and set $\underline{b} = (\frac{\alpha A}{2B})^{\frac{1}{2-\alpha}}$. Then the function $f(b) = Ab^\alpha - Bb^2$ satisfies the estimate*

$$\int_0^\infty \exp(f(b))db \le \exp(A\underline{b}^\alpha - B\underline{b}^2)\left(\underline{b} + \sqrt{\frac{2\pi}{\min(2B, 2B(2-\alpha))}}\right). \tag{3.14}$$

Proof. Note that

$$\begin{aligned} f'(b) &= \alpha A b^{\alpha-1} - 2Bb, \\ f''(b) &= \alpha(\alpha-1)Ab^{\alpha-2} - 2B, \\ f'''(b) &= \alpha(\alpha-1)(\alpha-2)Ab^{\alpha-3}. \end{aligned}$$

Let $\underline{b}$ be as given in statement of the lemma and note that $f'(b) = 0$ for $b = \underline{b}$, $f'(b) > 0$ for $b < \underline{b}$ and $f'(b) < 0$ for $b > \underline{b}$. Thus, $\underline{b}$ is the unique global maximum of $f(b)$. Upon inspecting the derivatives of f, it follows that $f''(b) < -2B < 0$ for $\alpha \le 1$, and $f'''(b) < 0$ for $1 < \alpha < 2$. This implies that for $b > \underline{b}$,

$$f''(b) < f''(\underline{b}) = -2B(2-\alpha),$$

and taking the Taylor expansion of f around $\underline{b}$ gives

$$f(b) = A\underline{b}^\alpha - B\underline{b}^2 + \frac{1}{2}(b-\underline{b})^2 f''(\xi(b))$$

for some $\xi(b) \in [b, \underline{b}]$ if $b < \underline{b}$ and $\xi(b) \in [\underline{b}, b]$ if $b > \underline{b}$. Note that for $b > \underline{b}$, $f''(\xi(b)) < \max(-2B, -2B(2-\alpha))$. Thus,

$$\begin{aligned} &\int_0^\infty \exp(Ab^\alpha - Bb^2)db \\ &\quad = \int_0^{\underline{b}} \exp(Ab^\alpha - Bb^2)db + \int_{\underline{b}}^\infty \exp(Ab^\alpha - Bb^2)db \\ &\quad \le \exp(A\underline{b}^\alpha - B\underline{b}^2)\left(\underline{b} + \int_{\underline{b}}^\infty \exp\left(-\frac{1}{2}(b-\underline{b})^2 \min(2B, 2B(2-\alpha))\right)db\right) \\ &\quad \le \exp(A\underline{b}^\alpha - B\underline{b}^2)\left(\underline{b} + \int_{-\infty}^\infty \exp\left(-\frac{(b-\underline{b})^2}{2(1/\min(2B, 2B(2-\alpha)))}\right)db\right) \\ &\quad = \exp(A\underline{b}^\alpha - B\underline{b}^2)\left(\underline{b} + \sqrt{\frac{2\pi}{\min(2B, 2B(2-\alpha))}}\right). \end{aligned}$$

□

Proof. (of Theorem 3.6) By Lemma 3.2.5, Lemma 3.2.4 can be applied to set $A = \mathbb{W}_T$, which implies (i). To prove (ii), set $M = \frac{1}{2}$ in Lemma 3.2.1 to prove the existence of a maximizer for 3.2. Analogously, $h \equiv 0$, $M = 1$ yield a maximizer for 3.3.
It remains to prove (iii). In view of (i), and since $\int_0^T (\dot{h}_t - \dot{x}_t)^2 dt \ge 0$, for any $h \in \mathbb{H}_T$ it follows that

$$\begin{aligned} L(h) &= \sup_{x \in \mathbb{H}_T}\left(2F(x) + \frac{1}{2}\int_0^T (\dot{x}_t - \dot{h}_t)^2 dt - \int_0^T \dot{x}_t^2 dt\right) \\ &\ge \sup_{x \in \mathbb{H}_T}\left(2F(x) - \int_0^T \dot{x}_t^2 dt\right), \end{aligned} \tag{3.15}$$

which implies the inequality

$$\inf_{h\in\mathbb{H}_T} L(h) \geq 2F(\hat{h}) - \int_0^T \dot{\hat{h}}_t^2 dt, \tag{3.16}$$

and hence $\hat{h}$ is asymptotically optimal if 3.4 is satisfied. For the uniqueness part, consider two distinct solutions h, g to 3.3. Strict convexity implies that

$$\begin{aligned} L(h) &\geq 2F(g) + \frac{1}{2}\int_0^T (\dot{g}_t - \dot{h}_t)^2 dt - \int_0^T \dot{g}_t^2 dt > 2F(g) - \int_0^T \dot{g}_t^2 dt \\ &= 2F(h) - \int_0^T \dot{h}_t^2 dt, \end{aligned}$$

which contradicts the optimality of h, and the uniqueness follows. □

3.3 Optimal change of drift for Asian options

This section employs Theorem 3.1.5 to find explicit formulas for the asymptotically optimal changes of drift, for geometric and arithmetic average Asian call and put options.

3.3.1 Optimal change of drift for the geometric average Asian call option[5]

Denoting by S_t the asset price at time t and by K the strike price, the payoff of a geometric average Asian call option is given by

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

Let $a = \sigma/T$ and $c = \frac{K}{S_0}\exp(-(r - \frac{\sigma^2}{2})\frac{T}{2})$. With

$$\begin{aligned} \left(e^{\frac{1}{T}\int_0^T \log S_t dt} - K\right)^+ &= \left(e^{\frac{1}{T}\int_0^T \log\left(S_0 e^{\sigma W_t + \left(r - \frac{\sigma^2}{2}\right)t}\right)dt} - K\right)^+ \\ &= \left(S_0 e^{\frac{1}{T}\int_0^T \sigma W_t dt + \frac{1}{T}\int_0^T \left(r - \frac{\sigma^2}{2}\right)t dt} - K\right)^+ \\ &= \left(S_0 e^{\frac{\sigma}{T}\int_0^T W_t dt + \left(r - \frac{\sigma^2}{2}\right)\frac{T}{2}} - K\right)^+ \\ &= K \underbrace{\frac{S_0}{K} e^{\left(r - \frac{\sigma^2}{2}\right)\frac{T}{2}}}_{1/c} \left(e^{\frac{\sigma}{T}\int_0^T W_t dt} - c\right)^+ \\ &= \frac{K}{c}\left(e^{a\int_0^T W_t dt} - c\right)^+ \end{aligned}$$

[5] cf. [4, example 4.1]

the payoff can be rewritten as functional $G : \mathbb{W}_T \to \mathbb{R}^+$, depending on x, the path of the Brownian motion of the asset price

$$G(x) = \frac{K}{c}\left(e^{a\int_0^T x_t dt} - c\right)^+. \tag{3.17}$$

Before we can use Theorem 3.1.5 we have to check Assumption 3.1.4. Therefore note that $F(x) = -\infty, \forall x$ with $G(x) = 0$, while on the set $G(x) > 0$, choosing $\alpha = 1,\ K_1 = \log\frac{K}{c}$ and $K_2 = aT$ gives $\log(\frac{K}{c}) + \log(e^{a\int_0^T x_t dt} - c) \leq \log(\frac{K}{c}) + aT \max_{t\in[0,T]} |x_t|$, which holds. So the requirement for the Theorem is satisfied.
Also note that condition 3.4 is certainly satisfied since F is concave. So if we find a solution for 3.3, this will be the unique and asymptotically optimal change of drift.

Now substitute 3.17 into 3.3, disregarding $2\log(\frac{K}{c})$, since it gives no contribution to finding the maximizer

$$\max_{x\in\mathbb{H}_T}\left(2\log(e^{a\int_0^T x_t dt} - c) - \int_0^T \dot{x}_t^2 dt\right). \tag{3.18}$$

Now we have to find the corresponding Euler-Lagrange equation. Therefore recall 2.28 and let

$$\begin{aligned}
&L_1(t,x,\dot{x}) = x,\\
&L_2(t,x,\dot{x}) = \dot{x}^2,\\
&J_1[x] = \int_0^T L_1(t,x,\dot{x})dt,\\
&J_2[x] = \int_0^T L_2(t,x,\dot{x})dt,\\
&F[J_1,J_2] = 2\log(e^{aJ_1} - c) - J_2.
\end{aligned}$$

With

$$\begin{aligned}
&\frac{\partial L_1}{\partial x} = 1,\\
&\frac{\partial L_1}{\partial \dot{x}} = 0,\\
&\frac{d}{dt}\frac{\partial L_1}{\partial \dot{x}} = 0\\
,&\frac{\partial L_2}{\partial x} = 0\\
&\frac{\partial L_2}{\partial \dot{x}} = 2\dot{x},\\
&\frac{d}{dt}\frac{\partial L_2}{\partial \dot{x}} = 2\ddot{x},
\end{aligned}$$

we get

$$-2\ddot{x}$$

on the left hand side, while with

$$\frac{\partial F}{\partial J_1} = \frac{2ae^{aJ_1}}{e^{aJ_1} - c},$$
$$\frac{\partial F}{\partial J_2} = -1,$$

we get

$$\frac{2ae^{aJ_1}}{e^{aJ_1} - c}$$

on the right hand side.

Thus we get the Euler-Lagrange ordinary differential equation

$$\ddot{x}_t = -\beta, \tag{3.19}$$

with

$$\beta = a\frac{e^{a\int_0^T x_t dt}}{e^{a\int_0^T x_t dt} - c}. \tag{3.20}$$

Note that [4, 4.4] states β by mistake without a in front of the integrals.
Hence, all solutions are of the form

$$x_t = -\frac{\beta}{2}t^2 + \gamma t \tag{3.21}$$

and therefore belong to $\mathbb{H}_T$.

Now we want to find an expression for γ just depending on β. Substituting 3.21 into 3.20 gives

$$\beta = a\frac{e^{-\frac{a\beta T^3}{6}+\frac{a\gamma T^2}{2}}}{e^{-\frac{a\beta T^3}{6}+\frac{a\gamma T^2}{2}} - c}.$$

Solving this equation for γ gives the wanted expression

$$\gamma(\beta) = \frac{a\beta T^3 - 6\log(\frac{\beta-a}{c\beta})}{3aT^2} \tag{3.22}$$

Now we try to find the maximum of 3.18. Therefore we substitute 3.21 into 3.18, which gives

$$2\log\left(e^{-\frac{aT^2}{6}(\beta T-3\gamma)} - c\right) - \frac{\beta^2T^3}{3} + \beta\gamma T^2 - \gamma^2 T.$$

Substituting 3.22 into the last result gives

$$-\frac{1}{9}\beta^2T^3 + 2\log\left(\frac{ac}{\beta - a}\right) + \frac{2\beta\log\left(\frac{\beta c}{\beta-a}\right)}{3a} - \frac{4\left(\log\left(\frac{\beta c}{\beta-a}\right)\right)^2}{a^2T^3},$$

what is well-defined for $\beta > a$.

Now taking the derivative with respect to β gives

$$\frac{2}{9aT^3}\left(\frac{-12a\beta T^3 - a\beta^2 T^6(\beta-a) - 3\beta T^3 \log\left(\frac{\beta-a}{c\beta}\right)(\beta-a) - 36\log\left(\frac{\beta-a}{c\beta}\right)}{\beta(\beta-a)}\right).$$

Let the result above be equal to zero. Then it reduces to

$$a\beta T^3 + 3\log\left(\frac{\beta-a}{c\beta}\right) = 0, \tag{3.23}$$

where solving for β over $\beta > a$ gives the optimal $\hat{\beta}$, which is unique by strict concavity of F.
Substituting 3.23 into 3.22 gives $\gamma(\hat{\beta}) = \hat{\beta}T$ and therefore the optimal change of drift for the geometric average Asian call option is given by

$$\hat{x}_t = -\frac{\hat{\beta}}{2}t^2 + \hat{\beta}Tt. \tag{3.24}$$

3.3.2 Optimal change of drift for the geometric average Asian put option

Now we want to find the asymptotically optimal change of drift for a geometric average Asian put option. This is done for the first time and can not be found in [4]. Consider S_t and K as before, then the payoff is given by

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

With $a = \sigma/T$ and $c = \frac{K}{S_0}\exp(-(r-\frac{\sigma^2}{2})\frac{T}{2})$, the payoff can be rewritten as functional $G : \mathbb{W}_T \to \mathbb{R}^+$, depending on x, the path of the Brownian motion of the asset price

$$G(x) = \frac{K}{c}\left(c - e^{a\int_0^T x_t dt}\right)^+. \tag{3.25}$$

To check Assumption 3.1.4. recall that $F(x) = -\infty, \forall x$ with $G(x) = 0$. On the set $G(x) > 0$, choosing $\alpha = 1$, $K_1 = \log\frac{K}{c} + \log(c)$ and $K_2 = 1$ gives $\log(\frac{K}{c}) + \log(c - e^{a\int_0^T x_t dt}) \leq \log(\frac{K}{c}) + \log(c) + \max\limits_{t\in[0,T]}|x_t|$, which holds. So the requirement for Theorem 3.1.5 is satisfied.
Since F is a reflection of the concave function of the case of the call option across a line through $\log(\frac{K}{c})$ parallel to the y-axis, it is again concave. So if we find a solution for 3.3, this will be the unique and asymptotically optimal change of drift, since condition 3.4 is certainly satisfied.

Substituting 3.25 into 3.3, disregarding $2\log(\frac{K}{c})$, since it gives no contribution to finding the maximizer, yields

$$\max_{x\in\mathbb{H}_T}\left(2\log(c - e^{a\int_0^T x_t dt}) - \int_0^T \dot{x}_t^2 dt\right). \tag{3.26}$$

Now we have to find the corresponding Euler-Lagrange equation. Since the only difference to the case of the call option is that $F[J_1, J_2] = 2\log(c - e^{aJ_1}) - J_2$ instead of $F[J_1, J_2] = 2\log(e^{aJ_1} - c) - J_2$, there is no change on the left hand side, but we have to take a look at the right hand side. With

$$\frac{\partial F}{\partial J_1} = -\frac{2ae^{aJ_1}}{c - e^{aJ_1}} = \frac{2ae^{aJ_1}}{e^{aJ_1} - c},$$
$$\frac{\partial F}{\partial J_2} = -1,$$

we get

$$\frac{2ae^{aJ_1}}{e^{aJ_1} - c}$$

on the right hand side. As one can see, also the right hand side did not change, therefore we get the same Euler-Lagrange ordinary differential equation

$$\ddot{x}_t = -\beta, \tag{3.27}$$

with the same β

$$\beta = a\frac{e^{a\int_0^T x_t dt}}{e^{a\int_0^T x_t dt - c}}. \tag{3.28}$$

Hence we get the same family of solutions

$$x_t = -\frac{\beta}{2}t^2 + \gamma t \tag{3.29}$$

and the same γ

$$\gamma(\beta) = \frac{a\beta T^3 - 6\log(\frac{\beta - a}{c\beta})}{3aT^2}. \tag{3.30}$$

Now we try to find the maximum of 3.26. Therefore we substitute 3.29 into 3.26, which gives

$$2\log\left(c - e^{-\frac{aT^2}{6}(\beta T - 3\gamma)}\right) - \frac{\beta^2 T^3}{3} + \beta\gamma T^2 - \gamma^2 T.$$

Substituting 3.30 into the last result yields

$$-\frac{1}{9}\beta^2 T^3 + 2\log\left(\frac{ac}{a-\beta}\right) + \frac{2\beta\log\left(\frac{\beta c}{\beta - a}\right)}{3a} - \frac{4\left(\log\left(\frac{\beta c}{\beta - a}\right)\right)^2}{a^2 T^3}.$$

Here one can see that this time $\beta < 0$ has to be satisfied. Now taking the derivative with respect to β states

$$\frac{2}{9aT^3}\left(\frac{-12a\beta T^3 - a\beta^2 T^6(\beta - a) - 3\beta T^3\log\left(\frac{\beta - a}{c\beta}\right)(\beta - a) - 36\log\left(\frac{\beta - a}{c\beta}\right)}{\beta(\beta - a)}\right),$$

which is identical to the result at this level in the case of the call option. Therefore it reduces to the same equation

$$a\beta T^3 + 3\log\left(\frac{\beta - a}{c\beta}\right) = 0, \tag{3.31}$$

where solving for β over $\beta < 0$ gives the optimal $\hat{\beta}$, which is unique by strict concavity of F.
Hence, by substituting 3.31 into 3.30, again we get $\gamma(\hat{\beta}) = \hat{\beta}T$ and the optimal change of drift for the geometric average Asian put option

$$\hat{x}_t = -\frac{\hat{\beta}}{2}t^2 + \hat{\beta}Tt. \tag{3.32}$$

As one can see, we mainly achieved the same results for calculating the asymptotically optimal change of drift for a geometric average Asian call and put option. Since the functionals 3.18 and 3.26, that have to be maximized, are not the same, one gets different optimal values $\hat{\beta}$ for those cases, even if all other parameters, like T, σ, S_0, K, r, would be the same. This means that the optimal change of drift for geometric average Asian call and put options belong to the same family of solutions 3.21, but in fact differs depending on the choice of the type of option(call,put) as well as of the choice of the parameters T, σ, S_0, K, r, since the optimal value $\hat{\beta}$ always changes.

3.3.3 Optimal change of drift for the arithmetic average Asian call option[6]

Denoting by S_t the asset price at time t and by K the strike price, the payoff of an arithmetic average Asian call option is

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

Let $a = \sigma$, $b = r - \frac{1}{2}\sigma^2$, $c = K\frac{T}{S_0}$ and $d = \frac{S_0}{T}$. With

$$\begin{aligned}\left(\frac{1}{T}\int_0^T S_t dt - K\right)^+ &= \left(\frac{1}{T}\int_0^T S_0 e^{\sigma W_t + \left(r - \frac{\sigma^2}{2}\right)t} dt - K\right)^+ \\ &= \frac{S_0}{T}\left(\int_0^T e^{\sigma W_t + \left(r - \frac{\sigma^2}{2}\right)t} dt - K\frac{T}{S_0}\right)^+\end{aligned}$$

the payoff can be rewritten as functional $G : \mathbb{W}_T \to \mathbb{R}^+$ depending on x, the path of the Brownian motion of the asset price

$$G(x) = d\left(\int_0^T e^{ax_t + bt} dt - c\right)^+. \tag{3.33}$$

To check Assumption 3.1.4., let $\alpha = 1, K_1 = \log d + \log(\frac{e^{bT}-1}{b}), K_2 = a$, which gives $\log(d) + \log(\int_0^T ax_t + btdt - c) \leq \log(d) + \log(\frac{e^{bT}-1}{b}) + a \max_{t\in[0,T]} |x_t|$, which holds. So the requirement for Theorem 3.1.5 is satisfied.
Now substitute 3.33 into 3.3, then we have

$$\max_{x\in\mathbb{H}_T}\left(2\log d + 2\log\left(\int_0^T e^{ax_t+bt}dt - c\right) - \int_0^T \dot{x}_t^2 dt\right). \tag{3.34}$$

[6] cf. [4, example 4.2]

Now we have to find the corresponding Euler-Lagrange equation. Therefore recall 2.28 and let

$$\begin{aligned}
&L_1(t, x, \dot{x}) = e^{ax+bt}, \\
&L_2(t, x, \dot{x}) = \dot{x}^2, \\
&J_1[x] = \int_0^T L_1(t, x, \dot{x}) dt, \\
&J_2[x] = \int_0^T L_2(t, x, \dot{x}) dt, \\
&F[J_1, J_2] = 2 \log(J_1 - c) - J_2.
\end{aligned}$$

With

$$\begin{aligned}
&\frac{\partial L_1}{\partial x} = a e^{ax+bt}, \\
&\frac{\partial L_1}{\partial \dot{x}} = 0, \\
&\frac{d}{dt} \frac{\partial L_1}{\partial \dot{x}} = 0 \\
&, \frac{\partial L_2}{\partial x} = 0 \\
&\frac{\partial L_2}{\partial \dot{x}} = 2\dot{x}, \\
&\frac{d}{dt} \frac{\partial L_2}{\partial \dot{x}} = 2\ddot{x},
\end{aligned}$$

we get

$$\frac{-2\ddot{x}}{a e^{ax+bt}}$$

on the left hand side, while with

$$\begin{aligned}
&\frac{\partial F}{\partial J_1} = \frac{2}{J_1 - c}, \\
&\frac{\partial F}{\partial J_2} = -1,
\end{aligned}$$

we get

$$\frac{2}{J_1 - c}$$

on the right hand side.

Thus we get the Euler-Lagrange ordinary differential equation

$$\ddot{x}_t = \lambda e^{a x_t + bt}, \tag{3.35}$$

with

$$\lambda = -\frac{a}{\int_0^T e^{a x_t + bt} dt - c}. \tag{3.36}$$

Equation 3.35 admits the family of solutions

$$x_t = \frac{\beta - b}{a}t - \frac{2}{a}\log\left(\frac{e^{\beta t}+\gamma}{1+\gamma}\right) \tag{3.37}$$

which belongs to $\mathbb{H}_T$.

Taking the second derivative of 3.37 with respect to t gives

$$\frac{\partial^2 x_t}{\partial t^2} = -\frac{2}{a}\frac{\beta^2\gamma e^{\beta t}}{(e^{\beta t}+\gamma)^2}.$$

Substituting this result and 3.37 into 3.35, we find out how the pair of parameters (β, γ) are linked to λ

$$\begin{aligned}
\ddot{x}_t &= \lambda e^{ax_t+bt} \\
\Leftrightarrow -\frac{2}{a}\frac{\beta^2\gamma e^{\beta t}}{(e^{\beta t}+\gamma)^2} &= \lambda e^{\beta t - 2\log(\frac{e^{\beta t}+\gamma}{1+\gamma})} \\
\Leftrightarrow -\frac{2}{a}\frac{\beta^2\gamma e^{\beta t}}{(e^{\beta t}+\gamma)^2} &= \lambda e^{\beta t}\frac{(1+\gamma)^2}{(e^{\beta t}+\gamma)^2} \\
\Leftrightarrow -\frac{2}{a}\beta^2\gamma &= \lambda(1+\gamma)^2 \\
\Leftrightarrow \lambda &= -\frac{2\beta^2\gamma}{a(1+\gamma)^2}
\end{aligned} \tag{3.38}$$

Note that in [4] it says that substituting (4.12) into (4.11) gives the above result, but (4.12) should be substituted into (4.10).
Since $\int_0^T e^{ax_t+bt}dt = \frac{(e^{\beta t}-1)(\gamma+1)}{\beta(e^{\beta T}+\gamma)}$, eliminating λ from 3.36 and 3.38 yields

$$\frac{a\beta(e^{\beta T}+\gamma)}{(\gamma+1)(e^{\beta T}-1) - \beta c(e^{\beta T}+\gamma)} = \frac{2\beta^2\gamma}{a(1+\gamma)^2}. \tag{3.39}$$

Note that in [4, 4.14] there should be a T instead of the t.

Now for fixed β, the above equation defines a cubic polynomial in γ, which yields three solutions $\gamma_1, \gamma_2, \gamma_3$, all of them depending on β. These solutions are explicit, but not what one would call nice expressions, therefore they are not presented here. Substituting each γ_i into 3.34, one can find the corresponding maximizing β_i, which serves to evaluate γ_i. Then choose $\hat{\beta}$ and $\hat{\gamma}$ to be the pair (β_i, γ_i), which gives the highest value of 3.34 under the condition that $\Re(\gamma) > 0$ and $|\Im(\gamma)| < 0.0000001$(numerically zero).
To check condition 3.4, maximize the functional

$$2\log d + 2\log\left(\int_0^T \exp\left(ax_t + bt\right)dt - c\right) + \frac{1}{2}\int_0^T \left(\dot{x}_t - \dot{\hat{x}}_t\right)^2 dt - \int_0^T \dot{x}_t^2 dt$$

over $x \in \mathbb{H}_T$. Again we have to find the corresponding Euler-Lagrange equation. Therefore recall 2.28 and let

$$
\begin{aligned}
L_1(t, x, \dot{x}) &= e^{ax+bt}, \\
L_2(t, x, \dot{x}) &= -\dot{x}^2 - 2\dot{x}\dot{\hat{x}} + \dot{\hat{x}}^2, \\
J_1[x] &= \int_0^T L_1(t, x, \dot{x})dt, \\
J_2[x] &= \int_0^T L_2(t, x, \dot{x})dt, \\
F[J_1, J_2] &= 2\log(J_1 - c) - \frac{1}{2}J_2.
\end{aligned}
$$

With

$$
\begin{aligned}
&\frac{\partial L_1}{\partial x} = ae^{ax+bt}, \\
&\frac{\partial L_1}{\partial \dot{x}} = 0, \\
&\frac{d}{dt}\frac{\partial L_1}{\partial \dot{x}} = 0 \\
,&\frac{\partial L_2}{\partial x} = 0 \\
&\frac{\partial L_2}{\partial \dot{x}} = -2\dot{x} - 2\dot{\hat{x}}, \\
&\frac{d}{dt}\frac{\partial L_2}{\partial \dot{x}} = -2\ddot{x} - 2\ddot{\hat{x}},
\end{aligned}
$$

we get

$$
\frac{2\ddot{x} + 2\ddot{\hat{x}}}{ae^{ax+bt}}
$$

on the left hand side, while with

$$
\begin{aligned}
\frac{\partial F}{\partial J_1} &= \frac{2}{J_1 - c}, \\
\frac{\partial F}{\partial J_2} &= \frac{1}{2},
\end{aligned}
$$

we get

$$
-\frac{4}{J_1 - c},
$$

on the right hand side.

Thus we get the Euler-Lagrange ordinary differential equation

$$
\ddot{x}_t = 2\lambda e^{ax_t+bt} - \ddot{\hat{x}}_t,
$$

where λ is defined as in 3.36. This ordinary differential equation does not admit an explicit solution, except in the trivial case $\lambda = 0$. However, a numerical integration of the Euler-Lagrange equation shows that 3.4 holds with several significant digits.

3.3.4 Optimal change of drift for the arithmetic average Asian put option

Now we want to find the asymptotically optimal change of drift for an arithmetic average Asian put option. Since this is done for the first time, it cannot be found in [4]. Consider S_t and K as before., then the payoff is given by

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

With $a = \sigma$, $b = r - \frac{1}{2}\sigma^2$, $c = K\frac{T}{S_0}$ and $d = \frac{S_0}{T}$ the payoff can be rewritten as functional $G : \mathbb{W}_T \to \mathbb{R}^+$ depending on x, the path of the Brownian motion of the asset price

$$G(x) = d\left(c - \int_0^T e^{ax_t+bt} dt\right)^+. \tag{3.40}$$

To check Assumption 3.1.4., let $\alpha = 1, K_1 = \log d + \log(c), K_2 = 1$, which gives $\log(d) + \log(c - \int_0^T ax_t + btdt) \leq \log(d) + \log(c) + \max_{t\in[0,T]} |x_t|$, which holds. So the requirement for Theorem 3.1.5 is satisfied.
Now substitute 3.40 into 3.3, then we have

$$\max_{x\in\mathbb{H}_T}\left(2\log d + 2\log\left(c - \int_0^T e^{ax_t+bt}dt\right) - \int_0^T \dot{x}_t^2 dt\right). \tag{3.41}$$

Now we have to find the corresponding Euler-Lagrange equation. Since the only difference to the case of the call option is that $F[J_1, J_2] = 2\log(c - J_1) - J_2$ instead of $F[J_1, J_2] = 2\log(J_1 - c) - J_2$, there is no change on the left hand side. With

$$\frac{\partial F}{\partial J_1} = -\frac{2}{c - J_1} = \frac{2}{J_1 - c},$$
$$\frac{\partial F}{\partial J_2} = -1,$$

we get

$$\frac{2}{J_1 - c}$$

on the right hand side, which is equal to the one in the case of the call option. Thus we get the same Euler-Lagrange ordinary differential equation

$$\ddot{x}_t = \lambda e^{ax_t+bt}, \tag{3.42}$$

with the same λ

$$\lambda = -\frac{a}{\int_0^T e^{ax_t+bt}dt - c}. \tag{3.43}$$

Hence the same family of solutions is admitted

$$x_t = \frac{\beta - b}{a}t - \frac{2}{a}\log\left(\frac{e^{\beta t} + \gamma}{1+\gamma}\right). \tag{3.44}$$

Substituting 3.44 and the second derivative of it into 3.42, shows that the parameters (β, γ) are linked to λ as before

$$\lambda = -\frac{2\beta^2\gamma}{a(1+\gamma)^2}. \tag{3.45}$$

Also eliminating λ from 3.43 and 3.45 yields an already known result

$$\frac{a\beta(e^{\beta T}+\gamma)}{(\gamma+1)(e^{\beta T}-1)-\beta c(e^{\beta T}+\gamma)} = \frac{2\beta^2\gamma}{a(1+\gamma)^2}, \tag{3.46}$$

which delivers $\gamma_1, \gamma_2, \gamma_3$ by solving the cubic polynomial for a fixed β. These solutions are explicit, but not what one would call nice expressions, therefore they are not presented here. Substituting each γ_i into 3.34, one can find the corresponding maximizing β_i, which serves to evaluate γ_i. Then choose $\hat{\beta}$ and $\hat{\gamma}$ to be the pair (β_i, γ_i), which gives the highest value of 3.34 under the condition that $\Re(\gamma) < 0$ and $|\Im(\gamma)| < 0.0000001$(numerically zero).
Also in this case we have to check condition 3.4, therefore maximize the functional

$$2\log d + 2\log(c - \int_0^T \exp(ax_t + bt)dt) + \frac{1}{2}\int_0^T (\dot{x}_t - \dot{\hat{x}}_t)^2 dt - \int_0^T \dot{x}_t^2 dt$$

over $x \in \mathbb{H}_T$. Hence we have to derive the corresponding Euler-Lagrange equation. Here the difference to the case of the call option is that $F[J_1, J_2] = 2\log(c - J_1) - \frac{1}{2}J_2$ instead of $F[J_1, J_2] = 2\log(J_1 - c) - \frac{1}{2}J_2$, which means that just the right hand side has to be looked at. With

$$\frac{\partial F}{\partial J_1} = -\frac{2}{c - J_1} = \frac{2}{J_1 - c},$$
$$\frac{\partial F}{\partial J_2} = \frac{1}{2},$$

we get

$$-\frac{4}{J_1 - c},$$

on the right hand side.

Thus we get the Euler-Lagrange equation

$$\ddot{x}_t = 2\lambda e^{ax_t + bt} - \ddot{\hat{x}}_t,$$

where λ is defined as in 3.43. So again we have the same equation as in the case of the call option. Therefore also in this case this ordinary differential equation does not admit an explicit solution, except in the trivial case $\lambda = 0$. However, a numerical integration of the Euler-Lagrange equation shows that 3.4 holds with several significant digits.

Also for arithmetic average Asian options we got mainly the same results for call options as well as for put options. What makes the difference is again the achieved pair of optimal values $(\hat{\beta}, \hat{\gamma})$.

Chapter 4

Different Monte Carlo estimators

In this chapter the different Monte Carlo estimators - the ordinary one, the one using importance sampling and the one using the method of control variates - for arithmetic average Asian call and put options will be derived.

4.1 Monte Carlo estimator without importance sampling

The price of an option is, as we know, equal to the discounted expectation of the payoff. Since the payoff of an arithmetic average Asian call respectively put option is

$$\left(\frac{1}{n}\sum_{i=1}^{n} S_{t_i} - K\right)^{+},$$
$$\left(K - \frac{1}{n}\sum_{i=1}^{n} S_{t_i}\right)^{+},$$

and the underlying asset in the Black-Scholes model under the risk-neutral measure is given by

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

we achieve the Monte Carlo estimator of the price of an arithmetic Asian call respectively put option

$$\frac{1}{N}\sum_{i=1}^{N} e^{-rt_n}\left(\frac{1}{n+1}\sum_{j=0}^{n} S_0 e^{\sigma x_{t_j}^{(i)} + \left(r - \frac{\sigma^2}{2}\right)t_j} - K\right)^{+},$$
$$\frac{1}{N}\sum_{i=1}^{N} e^{-rt_n}\left(K - \frac{1}{n+1}\sum_{j=0}^{n} S_0 e^{\sigma x_{t_j}^{(i)} + \left(r - \frac{\sigma^2}{2}\right)t_j}\right)^{+},$$

where K is the strike, T the maturity, N the number of sample paths, $n+1$ the number of trading days, $0 = t_0 < t_1 < \ldots < t_n = T$ the corresponding trading days and $x_{t_j}^{(i)},\ i = 1, \ldots, N,\ j = 0, \ldots, n$ the value of sample path i at time t_j.

4.2 Monte Carlo estimator with importance sampling

Recall that the expectation of a payoff G after using importance sampling is given(as in 3.6) by

$$\mathbb{E}_{\mathbb{P}}\left[G\right] = \mathbb{E}_{\mathbb{Q}^{\hat{h}}}\left[G\left(W^* + \hat{h}\right)e^{-\int_0^T \dot{\hat{h}}_t dW_t^* - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}\right],$$

where $\mathbb{P}$ denotes the risk-neutral measure in the Black-Scholes model, $\mathbb{Q}^{\hat{h}}$ the equivalent measure, leading to a change of drift from r to $r + \dot{\hat{h}}$, when changing the probability measure from $\mathbb{P}$ to $\mathbb{Q}^{\hat{h}}$, W^* a standard Brownian motion with respect to $\mathbb{Q}^{\hat{h}}$ and $\hat{h}$ the solution to 3.3.

Therefore we get the following Monte Carlo estimator for an arithmetic average Asian call respectively put option

$$\frac{1}{N}\sum_{i=1}^{N} e^{-rt_n}\left(\frac{1}{n+1}\sum_{j=0}^{n} S_0 e^{\sigma\left(x_{t_j}^{(i)} + \hat{h}(t_j)\right) + \left(r - \frac{\sigma^2}{2}\right)t_j} - K\right)^+ e^{-\int_0^T \dot{\hat{h}}_t dx_t^{(i)} - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}, \tag{4.1}$$

$$\frac{1}{N}\sum_{i=1}^{N} e^{-rt_n}\left(K - \frac{1}{n+1}\sum_{j=0}^{n} S_0 e^{\sigma\left(x_{t_j}^{(i)} + \hat{h}(t_j)\right) + \left(r - \frac{\sigma^2}{2}\right)t_j}\right)^+ e^{-\int_0^T \dot{\hat{h}}_t dx_t^{(i)} - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}. \tag{4.2}$$

4.2.1 Using the asymptotically optimal drift of a geometric average Asian option

In section 3.3.1 and 3.3.2 the optimal change of drift for a geometric average Asian option was computed as

$$\hat{h}_t = -\frac{\hat{\beta}}{2}t^2 + \hat{\beta}Tt, \tag{4.3}$$

where solving

$$a\beta T^3 + 3\log\left(\frac{\beta - a}{c\beta}\right) = 0, \tag{4.4}$$

numerically over $\beta > a$ respectively $\beta < 0$ gives $\hat{\beta}$. The first and second derivatives of $\hat{h}_t$ with respect to t are given by

$$\begin{aligned}\dot{\hat{h}}_t &= \hat{\beta}(T - t),\\ \ddot{\hat{h}}_t &= -\hat{\beta}.\end{aligned}$$

While the second integral of 4.1 respectively 4.2, $\int_0^T \dot{\hat{h}}_t^2$, can be computed explicitly, we try to simplify the first integral, $\int_0^T \dot{\hat{h}}_t dx_t^{(i)}$, by using Itô's formula.
Consider $g \in C^{1,2}$ with $g(t,x) = x\dot{\hat{h}}_t$ and note that a standard Brownian motion is an Itô process for $u = 0$ and $v = 1$. Then we get with Itô's formula

$$\begin{aligned} W_T\dot{\hat{h}}_T &= 0 + \int_0^T (W_t\ddot{\hat{h}}_t + 0 + 0)dt + \int_0^T \dot{\hat{h}}_t dW_t\\ \Leftrightarrow \int_0^T \dot{\hat{h}}_t dW_t &= W_T\dot{\hat{h}}_T - \int_0^T W_t\ddot{\hat{h}}_t dt. \end{aligned} \tag{4.5}$$

Substituting $\dot{\hat{h}}$ and $\ddot{\hat{h}}$ into 4.5 yields

$$\int_0^T \dot{\hat{h}}_t dW_t = W_T\beta(T-T) + \int_0^T \beta W_t dt = \beta \int_0^T W_t dt.$$

Hence we have

$$\int_0^T \dot{\hat{h}}_t dx_t^{(i)} = \beta \int_0^T x_t^{(i)} dt, \ i = 1, 2, \ldots, N$$

what can be approximated by

$$T\beta \frac{1}{n+1} \sum_{j=0}^{n} x_{t_j}^{(i)}, \ i = 1, 2, \ldots, N.$$

Substituting the results of this subsection into 4.1 respectively 4.2, we derive the following Monte Carlo estimator for an arithmetic average Asian call respectively put option using Importance Sampling and the asymptotically optimal change of drift of a geometric average Asian option

$$\frac{1}{N}\sum_{i=1}^{N} e^{-rt_n}\left(\frac{1}{n+1}\sum_{j=0}^{n} S_0 e^{\sigma\left(x_{t_j}^{(i)} - \frac{\hat{\beta}}{2}t_j^2 + \hat{\beta}Tt_j\right) + \left(r - \frac{\sigma^2}{2}\right)t_j} - K\right)^+ e^{-T\beta\frac{1}{n+1}\sum_{j=0}^{n} x_{t_j}^{(i)} - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}, \tag{4.6}$$

$$\frac{1}{N}\sum_{i=1}^{N} e^{-rt_n}\left(\frac{1}{n+1}\sum_{j=0}^{n} S_0 e^{\sigma\left(x_{t_j}^{(i)} - \frac{\hat{\beta}}{2}t_j^2 + \hat{\beta}Tt_j\right) + \left(r - \frac{\sigma^2}{2}\right)t_j} - K\right)^+ e^{-T\beta\frac{1}{n+1}\sum_{j=0}^{n} x_{t_j}^{(i)} - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}. \tag{4.7}$$

4.2.2 Using the asymptotically optimal drift of an arithmetic average Asian option

In section 3.3.3 and 3.3.4 the optimal change of drift for an arithmetic average Asian option was computed as

$$\hat{h}_t = \frac{\beta - b}{a}t - \frac{2}{a}\log\left(\frac{e^{\beta t} + \gamma}{1 + \gamma}\right), \tag{4.8}$$

where β and γ can be found as described in those sections.
The first and second derivatives of $\hat{h}_t$ with respect to t are given by

$$\dot{\hat{h}}_t = \frac{\beta - b}{a} - \frac{2}{a}\frac{\beta e^{\beta t}}{e^{\beta t} + \gamma},$$
$$\ddot{\hat{h}}_t = -\frac{2}{a}\frac{\beta^2 \gamma e^{\beta t}}{(e^{\beta t} + \gamma)^2}.$$

Again we can achieve a simplification of the first integral of 4.1respectively 4.2, $\int_0^T \dot{\hat{h}}_t dx_t^{(i)}$, by using Itô's formula. Therefore substitute $\dot{\hat{h}}$ and $\ddot{\hat{h}}$ into 4.5, which gives

$$\int_0^T \dot{\hat{h}}_t dW_t = W_T\left(\frac{\beta - b}{a} - \frac{2}{a}\frac{\beta e^{\beta T}}{e^{\beta T} + \gamma}\right) + \int_0^T \frac{2\beta^2\gamma e^{\beta t}}{a\left(e^{\beta t} + \gamma\right)^2} W_t dt.$$

Hence we have

$$\int_0^T \dot{\hat{h}}_t dx_t^{(i)} = x_T^{(i)} \left(\frac{\beta - b}{a} - \frac{2}{a} \frac{\beta e^{\beta T}}{e^{\beta T} + \gamma} \right) + \frac{2\beta^2\gamma}{a} \int_0^T \frac{e^{\beta t}}{\left(e^{\beta t} + \gamma\right)^2} x_t^{(i)} dt, \; i = 1, 2, \dots, N$$

where the integral can be approximated by

$$T \frac{2\beta^2\gamma}{a} \frac{1}{n+1} \sum_{j=0}^{n} \frac{e^{\beta t_j}}{(e^{\beta t_j} + \gamma)^2} x_{t_j}^{(i)}, \; i = 1, 2, \dots, N.$$

Substituting the results of this subsection into 4.1 respectively 4.2, we derive the following Monte Carlo estimator for an arithmetic average Asian call respectively put option using importance sampling and the asymptotically optimal change of drift of an arithmetic average Asian option

$$\begin{aligned} &\frac{1}{N} \sum_{i=1}^{N} e^{-rt_n} \left(\frac{1}{n+1} \sum_{j=0}^{n} S_0 e^{\sigma\left(x_{t_j}^{(i)} + \frac{\beta - b}{a} t_j - \frac{2}{a} \log\left(\frac{e^{\beta t_j} + \gamma}{1+\gamma}\right)\right) + \left(r - \frac{\sigma^2}{2}\right) t_j} - K \right)^+ \\ &\times e^{-x_T^{(i)}\left(\frac{\beta-b}{a} - \frac{2}{a}\frac{\beta e^{\beta T}}{e^{\beta T}+\gamma}\right) - T\frac{2\beta^2\gamma}{a}\frac{1}{n+1}\sum_{j=0}^{n} \frac{e^{\beta t_j}}{\left(e^{\beta t_j}+\gamma\right)^2} x_{t_j}^{(i)} - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}, \end{aligned} \tag{4.9}$$

$$\begin{aligned} &\frac{1}{N} \sum_{i=1}^{N} e^{-rt_n} \left(K - \frac{1}{n+1} \sum_{j=0}^{n} S_0 e^{\sigma\left(x_{t_j}^{(i)} + \frac{\beta - b}{a} t_j - \frac{2}{a} \log\left(\frac{e^{\beta t_j} + \gamma}{1+\gamma}\right)\right) + \left(r - \frac{\sigma^2}{2}\right) t_j} \right)^+ \\ &\times e^{-x_T^{(i)}\left(\frac{\beta-b}{a} - \frac{2}{a}\frac{\beta e^{\beta T}}{e^{\beta T}+\gamma}\right) - T\frac{2\beta^2\gamma}{a}\frac{1}{n+1}\sum_{j=0}^{n} \frac{e^{\beta t_j}}{\left(e^{\beta t_j}+\gamma\right)^2} x_{t_j}^{(i)} - \frac{1}{2}\int_0^T \dot{\hat{h}}_t^2 dt}. \end{aligned} \tag{4.10}$$

4.3 Monte Carlo estimator using the method of control variates

In section 2.4.2 we derived the control variate estimator 2.18, which is

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

with optimal coefficient b given by

$$\hat{b}_n = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}.$$

Let $X_i, \; i = 1, \dots, N$ be the price of a geometric average Asian call respectively put option according to the ith sample path $x^{(i)}$ and $Y_i, \; i = 1, \dots, N$ the price of an arithmetic average Asian call respectively put option according to the same sample path. Furthermore let $\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$ be the sample mean of all X_i and $\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i$ the sample mean of all Y_i. Let C_0^g respectively P_0^g denote the price of a geometric average Asian call respectively put option given by the closed form solution under the Black-Scholes modell as examined in section 2.6.

Then the estimator of the price of an arithmetic average Asian call respectively put option using the price of a geometric average Asian call respectively put option as control variate is denoted for both cases by

$$\bar{Y}\left(\hat{b}_N\right) = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\left(X_i - \mathbb{E}[X]\right)\right). \tag{4.11}$$

Chapter 5

Results

In this chapter numerical results for pricing arithmetic average Asian call and put options will be presented. The price will be determined by the Monte Carlo estimators, which were derived in chapter 4. Hence we do a usual Monte Carlo simulation under the risk-neutral drift, two using importance sampling and the asymptotically optimal drifts for Asian options of geometric and arithmetic average type, and in the end one more simulation using the method of control variates.

5.1 Arithmetic average Asian call option

At first consider an arithmetic average Asian call option. With parameters $T = 1$, $r = 5\%$, $\sigma = 25\%$, $S_0 = 50$ and $K = 70$ this option may be called way out-of-money.
Figure 5.1 shows the price paths of the asset in the absence of random shocks, i.e. S_t without a Brownian motion, under the different drifts. What one can see here is that the changed drifts really move the asset price into a region, where it makes the option valuable. Also one sees that the price path corresponding to the geometric drift is very similar to the path corresponding to the arithmetic one, although the closed form expression is much simpler in the geometric case.

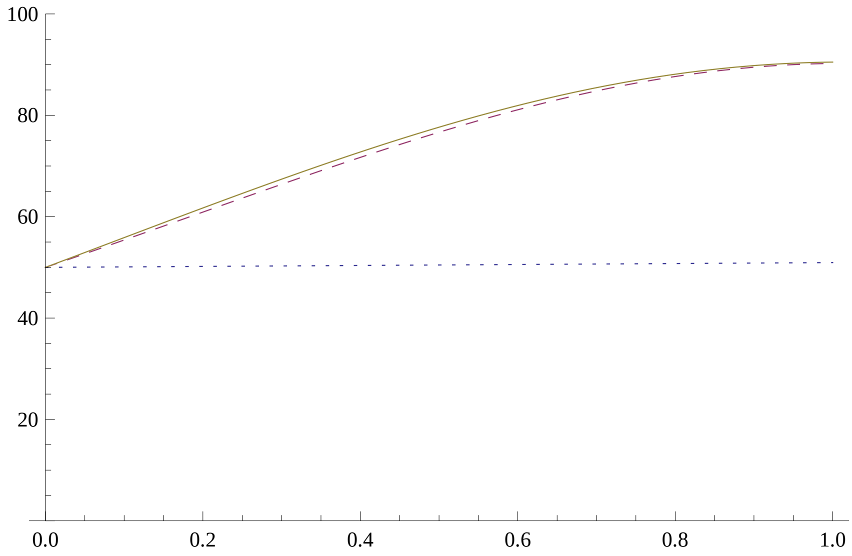

Figure 5.1: price path in absence of random shocks under risk-neutral drift(dotted line), geometric drift(solid line) and arithmetic drift(dashed line)

Figure 5.2 shows the value of 3.18, the functional that has to be maximized to find the optimal change of drift for a geometric average Asian call option, depending on β.

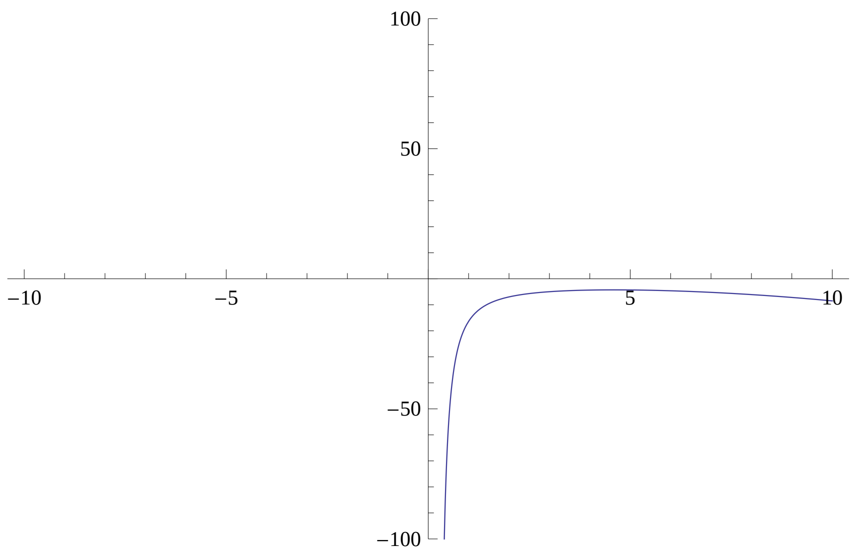

Figure 5.2: value of 3.18 depending on β

Figure 5.3 shows the value of 3.34, the functional that has to be maximized to find the optimal change of drift for an arithmetic average Asian call option, depending on β.

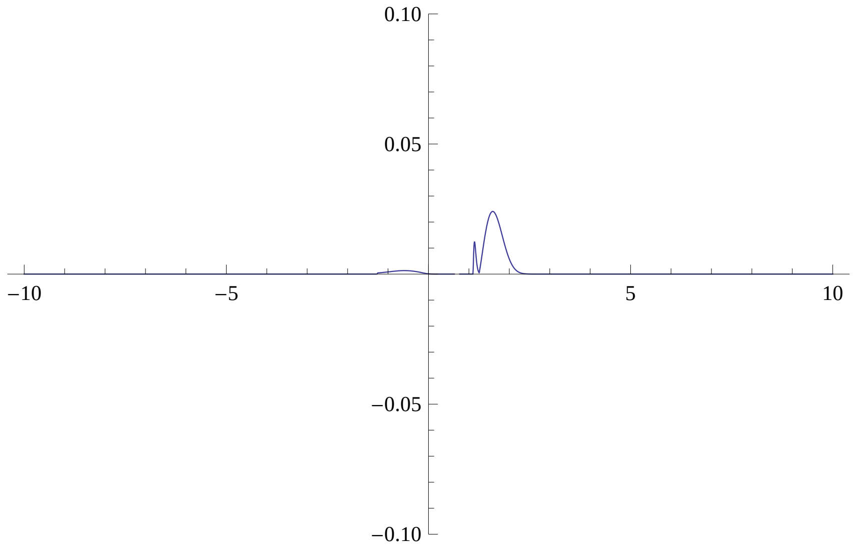

Figure 5.3: value of 3.34 depending on β

Table 5.1 compares the prices and standard errors according to the different Monte Carlo simulations and to different choices of sample size. Here one can see, that the estimator using the arithmetic drift has the lowest standard error, what seems logical, since the option is of arithmetic type. The estimator using the geometric drift follows immediately after it, also for smaller sample sizes, but it offers a much simpler alternative. The estimator using the method of control variates achieves a slightly higher standard error, but still a much better one than the usual estimator.

Sample size	Risk-neutral	Geometric drift	Arithmetic drift	Control variate
100000	5.92	5.74	5.74	5.78
	(0,212)	(0.019)	(0.018)	(0.033)
20000	6.16	5.73	5.74	5.82
	(0.472)	(0.041)	(0.041)	(0.074)
5000	6.51	5.72	5.73	5.90
	(0.898)	(0.084)	(0.082)	(0.141)

Table 5.1: Monte Carlo estimators of an arithmetic average Asian call option price using different drifts. Prices are in cents. Parameter values $T = 1$, $r = 5\%$, $\sigma = 25\%$, $S_0 = 50$, $K = 70$. Simulations are performed with a time-increment of 1/252, corresponding to one business day.

Table 5.2 now compares the performance of the estimators in terms of variance reduction over different strikes and volatilities. Therefore variance ratios, i.e. the variance of the usual Monte Carlo estimator divided by the variance of the others, are given. Additionally the corresponding prices are stated. As one can see the achieved variance reduction for the estimators using importance sampling increases with the strike and decreases with volatility. This means that the more unlikely it is for the option to be valuable, the more efficient are the estimators using importance sampling. In the case of the estimator using the method of control variates it is the other way around, the more likely it is for the option to be valuable, the more efficient is the estimator. So for pricing way out-of-money average Asian call options, importance sampling seems to be the better choice. Since the option is of arithmetic type, the estimator using the arithmetic drift achieves more reduction of variance than the one using the geometric drift.

Volatility	Strike	Price(Variance Ratio) geometric	arithmetic	control variate
10%	50	182.19(7.09)	182.18(7.12)	181.98(4692)
	60	0.37(486)	0.37(487)	0.37(71)
15%	50	234.44(7.71)	234.43(7.77)	234.16(2255)
	60	7.03(64.21)	7.03(64.24)	7.04(146)
20%	50	288.29(8.26)	288.26(8.33)	287.96(1305)
	60	26.31(32.44)	26.30(32.51)	26.26(204)
	70	0.98(458)	0.98(470)	0.98(28.4)
25%	50	342.73(8.78)	342.69(8.88)	342.35(846)
	60	56.79(23.66)	56.78(23.77)	56.67(218)
	70	5.74(131)	5.74(134)	5.78(40)
30%	50	397.41(9.30)	397.37(9.42)	396.99(588)
	60	95.26(20.06)	95.24(20.23)	95.10(206)
	70	16.99(69)	16.99(70)	17.08(53)
	80	2.54(337)	2.55(356)	2.53(19.3)
35%	50	452.16(9.82)	452.10(9.99)	451.71(430)
	60	139.14(18.31)	139.11(18.53)	138.88(189)
	70	35.61(47)	35.61(48)	35.58(64)
	80	8.21(151)	8.21(159)	8.27(23.9)
	90	1.80(578)	1.80(637)	1.79(13)
40%	50	506.90(10.37)	506.82(10.57)	506.42(327)
	60	186.76(17.4)	186.72(17.7)	186.44(168)
	70	61.17(37.4)	61.15(38.3)	61.07(69.2)
	80	18.82(92.2)	18.82(96.6)	19.03(28.3)
	90	5.64(254)	5.64(277)	5.59(15.8)

Table 5.2: Prices and variance ratios of an arithmetic average Asian call option over different volatilities and strikes corresponding to the different Monte Carlo estimators over 100000 sample paths. Prices are in cents. Parameter values $T = 1$, $r = 5\%$, $S_0 = 50$. Simulations are performed with a time-increment of 1/252, corresponding to one business day.

Table 5.3 states the optimal value $\hat{\beta}$geo for the optimal change of drift of a call option of geometric type as well as the optimal values $(\hat{\beta}, \hat{\gamma})$ for the optimal change of drift for a call option of arithmetic type corresponding to different values of K and σ.

Volatility	Strike	$\hat{\beta}$geo	$\hat{\beta}$	$\hat{\gamma}$
10%	50	1.45743	0.5515	2.04397
	60	5.35967	1.0635	3.15267
15%	50	1.59187	0.7075	2.26444
	60	4.01956	-1.1325	0.301007
20%	50	1.67831	0.8415	2.49156
	60	3.41506	1.2095	3.52442
	70	5.38933	-1.5345	0.207195
25%	50	1.74409	-0.9625	0.367644
	60	3.08833	1.2915	3.74261
	70	4.59629	1.5895	5.02015
30%	50	1.79915	-1.0745	0.338734
	60	2.893	1.3735	3,98098
	70	4.09965	-1.6505	0.190965
	80	5.26205	-1.8865	0.150906
35%	50	1.84799	1.1795	3.19304
	60	2.76917	-1.4565	0.236498
	70	3.7679	-1.7145	0.182575
	80	4.73503	1.9375	6.86671
	90	5.63593	2.1315	8.3396
40%	50	1.89291	-1.2805	0.291363
	60	2.68821	1.5395	4.4847
	70	3.53632	-1.7805	0.174208
	80	4.35935	1.9935	7.12413
	90	5.12978	2.1795	8.60169

Table 5.3: Optimal values for the optimal changes of drifts for call options of geometric($\hat{\beta}$geo) and arithmetic($\hat{\beta}$,$\hat{\gamma}$) type corresponding to different values of K and σ. Parameter values $T = 1$, $r = 5\%$, $S_0 = 50$.

5.2 Arithmetic average Asian put option

Now consider an arithmetic average Asian put option with parameters $T = 1$, $r = 5\%$, $\sigma = 25\%$, $S_0 = 50$ and $K = 35$.
Figure 5.4 shows the price paths of the asset in the absence of random shocks under the different drifts. This time the gap between the price paths under the geometric and under the arithmetic drift is not that small as in the case of the call option, but again the changed drifts move the asset price into a region, where it makes the option valuable. Since the price under the arithmetic drift is lower, it is more likely for the option to be valuable and will lead to a higher price of the option as under the geometric drift.

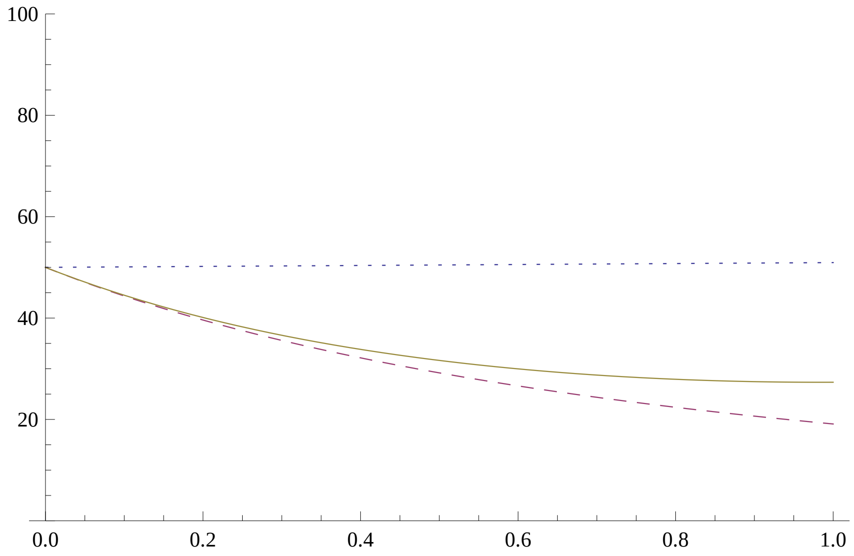

Figure 5.4: price path in absence of random shocks under risk-neutral drift(dotted line), geometric drift(solid line) and arithmetic drift(dashed line)

Figure 5.5 shows the value of 3.26, the functional that has to be maximized to find the optimal change of drift for a geometric average Asian put option, depending on β.

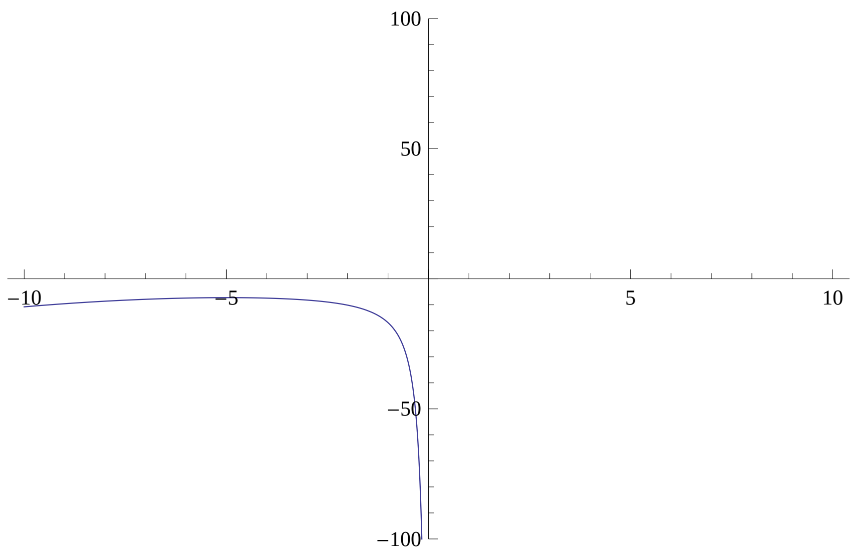

Figure 5.5: value of 3.26 depending on β

Figure 5.6 shows the value of 3.41, the functional that has to be maximized to find the optimal change of drift for an arithmetic average Asian put option, depending on β.

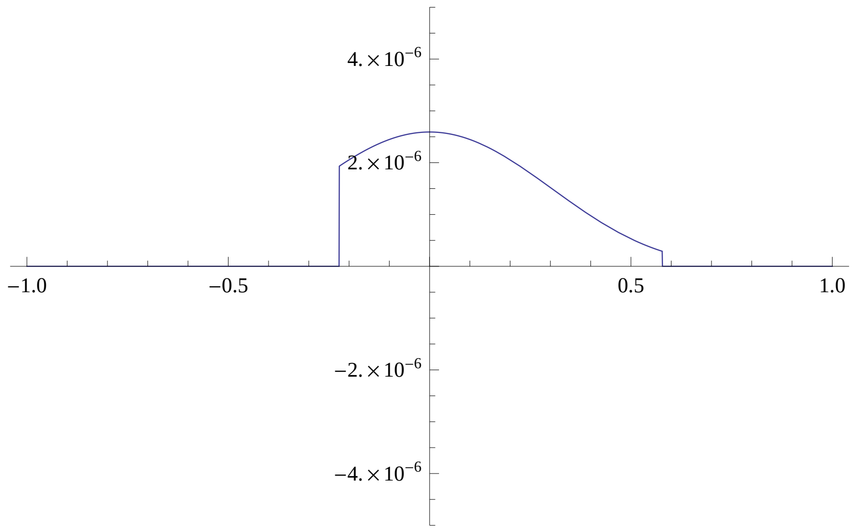

Figure 5.6: value of 3.41 depending on β

Table 5.4 compares the prices and standard errors according to the different Monte Carlo simulations and to different choices of sample size. Surprisingly this time the estimator using importance sampling and the geometric drift achieves the lowest standard error, followed by the estimator using the method of control variates. Also the estimator using importance sampling and the arithmetic drift has a lower standard error than the usual estimator, but this method is not as effective in reducing variance as before.

Sample size	Risk-neutral	Geometric drift	Arithmetic drift	Control variate
100000	0.444	0.519	0.528	0.510
	(0,0314)	(0.0020)	(0.0098)	(0.0075)
20000	0.475	0.519	0.513	0.505
	(0.0721)	(0.0044)	(0.0169)	(0.0177)
5000	0.346	0.521	0.494	0.422
	(0.1064)	(0.0088)	(0.0314)	(0.0356)

Table 5.4: Monte Carlo estimators of an arithmetic average Asian put option price using different drifts. Prices are in cents. Parameter values $T = 1$, $r = 5\%$, $\sigma = 25\%$, $S_0 = 50$, $K = 35$. Simulations are performed with a time-increment of 1/252, corresponding to one business day.

Table 5.5 compares the performance of the estimators in terms of variance reduction over different strikes and volatilities. Therefore again we take a look at variance ratios and the corresponding prices. In this situation one can see that the achieved variance reduction for the estimators using importance sampling decreases with the strike and with volatility. This means that, as in the case of the call option, the more unlikely it is for the option to be valuable, the more efficient are the estimators using importance sampling. In the case of the estimator using the method of control variates it is the other way around, the more likely it is for the option to be valuable, the more efficient is the estimator. So for pricing way out-of-money average Asian put options, importance sampling seems to be the better choice, where the geometric drift achieves a way better result corresponding to variance reduction. Also one can see, that some options can not be priced by the usual estimator and therefore also not by the one using the method of control variates, while the estimators using importance sampling achieve a result different from zero. This result is way smaller than 1 cent and therefore not really helpful in practice, but it emphasizes the advantage of the estimators using importance sampling.

Volatility	Strike	Price(Variance Ratio)		
		geometric	arithmetic	control variate
10%	50	61,04(8.81)	60,85(0.02)	61,07(6544)
	45	1,02(112)	1,07(0.25)	1,02(364)
	40	0,00029(2443)	0,00038(0.43)	0,00007(2.11)
15%	50	113,23(7.67)	114,78(0.07)	113,26(2907)
	45	11,01(27.4)	11,17(0.35)	11,00(462)
	40	0,181(474)	0,191(2.04)	0,179(39.6)
20%	50	167,02(7.05)	168,10(0.17)	167,07(1628)
	45	33,16(15.9)	33,47(0.51)	33,16(455)
	40	2,41(83.8)	2,48(1.80)	2,42(77.6)
	35	0,0351(2063)	0,0365(25.7)	0,0382(16.3)
	30	0,0000378(-)	0,0000379(-)	0(-)
25%	50	221,41(6.6)	222,33(0.31)	221,47(1035)
	45	64,15(11.86)	64,69(0.69)	64,15(389)
	40	9,68(35.6)	9,82(1.92)	9,65(93.0)
	35	0,5186(253)	0,5276(10.2)	0,5104(17.5)
	30	0,005219(11980)	0,005286(347)	0,006577(7.36)
30%	50	276,04(6.27)	277,03(0.47)	276,13(716)
	45	100,88(9.83)	101,64(0.88)	100,91(324)
	40	23,16(21.9)	23,39(2.02)	23,11(100.9)
	35	2,57(87.8)	2,59(7.21)	2,58(27.5)
	30	0,0885(961)	0,0899(55.8)	0,0907(6.57)
	25	0,00042(5933)	0,00043(235.6)	0,00016(2.00)
35%	50	330,75(5.98)	332,01(0.63)	330,86(524)
	45	141,34(8.60)	142,26(1.07)	141,39(272)
	40	42,47(15.9)	42,78(2.13)	42,45(103.0)
	35	7,41(44.6)	7,46(5.81)	7,38(31.9)
	30	0,5421(245)	0,5473(24.9)	0,5295(8.43)
	25	0,009203(6575)	0,009253(516)	0,012762(4.40)
	20	0,0000111(-)	0,0000112(-)	0(-)
40%	50	385,45(5.72)	386,91(0.83)	385,58(400)
	45	184,26(7.75)	185,27(1.29)	184,32(227)
	40	66,65(12.72)	67,03(2.31)	66,61(99.2)
	35	15,75(28.3)	15,84(4.98)	15,68(35.2)
	30	1,895(106)	1,908(16.0)	1,897(12.0)
	25	0,07359(1026)	0,07387(136)	0,07828(3.95)
	20	0,000370(4474)	0,000372(514)	0,000155(1.87)
	15	0,0000000283(-)	0,0000000287(-)	0(-)

Table 5.5: Prices and variance ratios of an arithmetic average Asian put option over different volatilities and strikes corresponding to the different Monte Carlo estimators over 100000 sample paths. Prices are in cents. Parameter values $T = 1$, $r = 5\%$, $S_0 = 50$. Simulations are performed with a time-increment of $1/252$, corresponding to one business day.

Table 5.6 states the optimal value $\hat{\beta}$geo for the optimal change of drift of a put option of geometric type as well as the optimal values $(\hat{\beta}, \hat{\gamma})$ for the optimal change of drift for a put option of arithmetic type corresponding to different values of K and σ.

Volatility	Strike	$\hat{\beta}$geo	$\hat{\beta}$	$\hat{\gamma}$
10%	50	-3.59663	-0.0005	-1.00102
	45	-4.49579	-0.0005	-1.00213
	40	-7.75374	-0.0005	-1.00153
15%	50	-1.90408	-0.0005	-1.00208
	45	-3.36655	-0.0005	-1.00169
	40	-5.3985	-0.0005	-1.00131
20%	50	-1.80295	-0.0005	-1.00169
	45	-2.82978	-0.0005	-1.00142
	40	-4.2603	-0.0005	-1.00114
	35	-6.06202	-0.0005	-1.00088
	30	-8.24682	-0.0005	-1.00065
25%	50	-1.73118	-0.0005	-1.00144
	45	-2.51433	-0.0005	-1.00123
	40	-3.59663	-0.0005	-1.00102
	35	-4.98034	-0.0005	-1.00081
	30	-6.68311	-0.0005	-1.00062
30%	50	-1.67372	-0.0005	-1.00126
	45	-2.30308	-0.0005	-1.00109
	40	-3.16267	-0.0005	-1.00092
	35	-4.27064	-0.0005	-1.00075
	30	-5.65056	-0.0005	-1.00058
	25	-7.35619	-0.0005	-1.00044
35%	50	-1.62435	-0.0005	-1.00112
	45	-2.14849	-0.0005	-1.00098
	40	-2.85549	-0.0005	-1.00084
	35	-3.76996	-0.0005	-1.0007
	30	-4.9194	-0.0005	-1.00055
	25	-6.35273	-0.0005	-1.00042
	20	-8.16545	-0.0005	-1.0003
40%	50	-1.58002	-0.0005	-1.00102
	45	-2.0279	-0.0005	-1.0009
	40	-2.62486	-0.0005	-1.00078
	35	-3.39737	-0.0005	-1.00065
	30	-4.37487	-0.0005	-1.00052
	25	-5.60326	-0.0005	-1.00041
	20	-7.167	-0.0005	-1.0003
	15	-9.2353	-0.0005	-1.0002

Table 5.6: Optimal values for the optimal changes of drifts for put options of geometric($\hat{\beta}$geo) and arithmetic($\hat{\beta}$,$\hat{\gamma}$) type corresponding to different values of K and σ. Parameter values $T = 1$, $r = 5\%$, $S_0 = 50$.

Appendix

Maple codes

To compute the results stated in the last chapter Maple 15 was used. In this section the corresponding codes are given.

```
> ####################################################################################
> ########################### arithmetic average Asian call option ########################
> ####################################################################################
> restart;
> with(Finance) :
> with(Statistics) :
>
> #################################### initialize values ###################################
> W := WienerProcess( ) :
> T := 1 :
> S := 50 :
> sigma := 0.25 :
> r := 0.05 :
> K := 70 :
> N := 100000 :
> N1 := 20000 :
> N2 := 5000 :
> n = 252 :
> ######################### N sample paths X of a Brownian motion ######################
> X := SamplePath(W(t), t = 0 ..1, timesteps = 252, replications = N) :
>
> ########################### geometric and arithmetic average ##########################
```

$$SGeo := Vector\left(\left[seq\left(S\cdot\left(mul\left(e^{\frac{\left(r-\frac{\sigma^2}{2}\right)\cdot(j-1)}{252}+\mathrm{sigma}\cdot X[i,j]}, j=1..253\right)\right)^{\frac{1}{253}}, i=1..N\right)\right]\right) :$$

$$SArith := Vector\left(\left[seq\left(\frac{S\cdot add\left(e^{\frac{\left(r-\frac{\sigma^2}{2}\right)\cdot(j-1)}{252}+\mathrm{sigma}\cdot X[i,j]}, j=1..253\right)}{253}, i=1..N\right)\right]\right) :$$

```
>
> ####################################################################################
> ############################## usual Monte Carlo estimator ############################
> ####################################################################################
>
> ########### vector containing N option prices corresponding to N sample paths ###########
```

$$Cneutral := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max(\mathrm{SArith}[i]-K, 0), i=1..N\right)\right]\right) :$$

```
>
> ################### price and standard error for N=100.000 in Cents ##################
> 100· Mean(Cneutral);
                                5.924380005                                          (1)
> 100· StandardError(Mean, Cneutral);
                                0.211959047419500                                    (2)
> ################### price and standard error for N=20.000 in Cents ###################
> C1neutral := Cneutral[1 ..20000] :
```

> $100 \cdot Mean(C1neutral);$

$$6.160881101 \tag{3}$$

> $100 \cdot StandardError(Mean, C1neutral);$

$$0.471899710961524 \tag{4}$$

> ##################### price and standard error for N=5.000 in Cents ####################

> $C2neutral := Cneutral[1..5000]:$

> $100 \cdot Mean(C2neutral);$

$$6.510154867 \tag{5}$$

> $100 \cdot StandardError(Mean, C2neutral);$

$$0.898334155279336 \tag{6}$$

> ##

> ######### Monte Carlo estimator using Importance Sampling and the geometic drift #########

> ##

> $aGeo := \frac{\text{sigma}}{T}:$

> $cGeo := \frac{K}{S} \cdot \mathrm{e}^{-\frac{\left(r - \frac{\sigma^2}{2}\right) \cdot T}{2}}:$

>

> ##################################### compute beta ####################################

> $equ := aGeo \cdot b \cdot T^3 + 3 \cdot \ln\left(\frac{(b - aGeo)}{cGeo \cdot b}\right) = 0:$

> $betaGeo := fsolve(equ, b, b = aGeo..10000);$

$$betaGeo := 4.596289154 \tag{7}$$

> ########################## aymptotically optimal change of drift #########################

> $h := t \to -\frac{betaGeo}{2} \cdot t^2 + betaGeo \cdot T \cdot t:$

> $h1 := t \to betaGeo \cdot (T - t):$

>

> ############################ computation of the rescaling terms ###########################

> $Rescale1 := Vector\left(\left[seq\left(betaGeo \cdot \frac{add(X[i,j], j = 1..253)}{253}, i = 1..N\right)\right]\right):$

> $Rescale2 := \mathrm{e}^{-0.5 \cdot \int_0^T (h1(t))^2 \, \mathrm{d}t}:$

> ########### vector containing N option prices corresponding to N sample paths ############

> $CGeo := Vector\left(\left[seq\left(\mathrm{e}^{-r \cdot T} \cdot \max\left(\frac{add\left(S \cdot \mathrm{e}^{\frac{\left(r - \frac{\sigma^2}{2}\right) \cdot (j-1)}{252} + \text{sigma} \cdot \left(h\left(\frac{(j-1)}{252}\right) + X[i,j]\right)}, j = 1..253\right)}{253} - K, 0\right) \cdot \mathrm{e}^{-Rescale1[i]}\right.\right.\right.$

$$\cdot Rescale2,\ i = 1..N\Bigg)\Bigg]\Bigg]\Bigg) :$$

> ################### *price and standard error for N=100.000 in Cents* ###################

> $100 \cdot Mean(CGeo);$

$$5.737148757 \qquad (8)$$

> $100 \cdot StandardError(Mean, CGeo);$

$$0.0185039720588016 \qquad (9)$$

> ################### *price and standard error for N=20.000 in Cents* ###################

> $C1Geo := CGeo[1..20000]:$

> $100 \cdot Mean(C1Geo);$

$$5.734282462 \qquad (10)$$

> $100 \cdot StandardError(Mean, C1Geo);$

$$0.0414330718958574 \qquad (11)$$

> #################### *price and standard error for N=5.000 in Cents* ###################

> $C2Geo := CGeo[1..5000]:$

> $100 \cdot Mean(C2Geo);$

$$5.720578536 \qquad (12)$$

> $100 \cdot StandardError(Mean, C2Geo);$

$$0.0836469739471332 \qquad (13)$$

>

> ##

> ######### *Monte Carlo estimator using Importance Sampling and the arithmetic drift* #########

> ##

> $a := \text{sigma}:$

> $b := r - \frac{\sigma^2}{2}:$

> $c := \frac{K \cdot T}{S}:$

> $d := \frac{S}{T}:$

> $\text{beta} := 1.5895;$

$$\beta := 1.5895 \qquad (14)$$

> $\text{delta} := 5.02015;$

$$\delta := 5.02015 \qquad (15)$$

>

> ######################### *aymptotically optimal change of drift* ########################

> $h := t \rightarrow \frac{\text{beta} - b}{a} \cdot t - \frac{2}{a} \cdot \ln\left(\frac{e^{\text{beta} \cdot t} + \text{delta}}{1 + \text{delta}}\right):$

> $h1 := t \rightarrow \frac{\text{beta} - b}{a} - \frac{2}{a} \cdot \frac{\text{beta} \cdot e^{\text{beta} \cdot t}}{e^{\text{beta} \cdot t} + \text{delta}}:$

>

> ########################## *computation of the rescaling terms* ##########################

$$> Rescale1 := Vector\left(\left[seq\left(h1(T)\cdot X[i,253] + T\cdot\frac{2\cdot\beta^2\cdot\delta}{a}\cdot\frac{add\left(\frac{e^{\frac{\mathrm{beta}\cdot(j-1)}{252}}}{\left(e^{\frac{\mathrm{beta}\cdot(j-1)}{252}}+\mathrm{delta}\right)^2}X[i,j], j=1..253\right)}{253}, i=1..N\right)\right]\right):$$

$$> Rescale2 := e^{-0.5\cdot\int_0^T (h1(t))^2\,dt}:$$

> ########### *vector containing N option prices corresponding to N sample paths* ############

$$> CArith := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max\left(\frac{add\left(S\cdot e^{\frac{\left(r-\frac{\sigma^2}{2}\right)\cdot(j-1)}{252}+\mathrm{sigma}\cdot\left(h\left(\frac{(j-1)}{252}\right)+X[i,j]\right)}, j=1..253\right)}{253}-K, 0\right)\cdot e^{-Rescale1[i]}\cdot Rescale2, i=1..N\right)\right]\right):$$

> #################### *price and standard error for N=100.000 in Cents* ###################

> $100\cdot Mean(CArith);$

5.738688869 **(16)**

> $100\cdot StandardError(Mean, CArith);$

0.0183120502045484 **(17)**

> #################### *price and standard error for N=20.000 in Cents* ####################

> $C1Arith := CArith[1..20000]:$

> $100\cdot Mean(C1Arith);$

5.743829812 **(18)**

> $100\cdot StandardError(Mean, C1Arith);$

0.0409914097726682 **(19)**

> ##################### *price and standard error for N=5.000 in Cents* ####################

> $C2Arith := CArith[1..5000]:$

> $100\cdot Mean(C2Arith);$

5.734950026 **(20)**

> $100\cdot StandardError(Mean, C2Arith);$

0.0824966285807213 **(21)**

>

```
> ###############################################################################
> ############### Monte Carlo estimator using the method of control variates ##############
> ###############################################################################
```

> $m := 252 + 1 :$

> $Digits := 20 :$

> $z := \text{sigma}\cdot\sqrt{\dfrac{2\cdot m+1}{6\cdot(m+1)}} :$

> $\text{rho} := \dfrac{\left(r - \dfrac{\sigma^2}{2} + z^2\right)}{2} :$

> $\text{phi} := RandomVariable(Normal(0, 1)) :$

>

> $d1 := \dfrac{\ln\left(\dfrac{S}{K}\right) + \left(\text{rho} + \dfrac{z^2}{2}\right)\cdot T}{\sqrt{T}\cdot z} :$

> $d2 := \dfrac{\ln\left(\dfrac{S}{K}\right) + \left(\text{rho}- \dfrac{z^2}{2}\right)\cdot T}{\sqrt{T}\cdot z} :$

>

```
> ############## computing Xi , Yi , Xbar, Ybar, EX as described in the method ##############
```

> $\text{Xi} := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max(\text{SGeo}[i]-K, 0), i = 1 ..N\right)\right]\right) :$

> $Yi := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max(\text{SArith}[i]-K, 0), i = 1 ..N\right)\right]\right) :$

> $Xbar := Mean(\text{Xi})$

$$Xbar := 0.041540792658249837363 \quad (22)$$

> $Ybar := Mean(Yi)$

$$Ybar := 0.059243800049585760919 \quad (23)$$

> $EX := e^{(\text{rho} - r)\cdot T}\cdot\left(S\cdot CDF(\text{phi}, d1) - K\cdot e^{-\text{rho}\cdot T}\cdot CDF(\text{phi}, d2)\right) :$

> $simplify(EX)$

$$0.040385618906451286000 \quad (24)$$

> ######### *optimal coefficient* b_N , *price* **and** *standard* **error for** $N = 100.000$ **in** *Cents* ########

> $sum1 := \text{add}((\text{Xi}[i] - Xbar)\cdot(\text{Yi}[i] - Ybar), i = 1 ..N) :$

> $\text{sum2} := \text{add}\left((\text{Xbar} - \text{Xi}[i])^2, i = 1 ..N\right) :$

> $b := \dfrac{sum1}{sum2};$

$$b := 1.2415352547355586260 \quad (25)$$

> $Ycv := Vector([seq(Yi[i] - b\cdot(\text{Xi}[i] - EX), i = 1 ..N)]) :$

> $100\cdot Mean(Ycv)$

$$5.7809611111382723646 \quad (26)$$

> $100\cdot StandardError(Mean, Ycv);$

$$0.033465479872684520483 \quad (27)$$

> ######### *optimal coefficient* b_{N1} , *price* **and** *standard* **error for** $N = 20.000$ **in** *Cents* #########

> $Xi1 := \mathrm{Xi}[1..20000]$:

> $Yi1 := Yi[1..20000]$:

> $Xbar1 := Mean(Xi1)$;

$$Xbar1 := 0.04313612465531437367 2 \tag{28}$$

> $Ybar1 := Mean(Yi1)$;

$$Ybar1 := 0.06160881100942021549 0 \tag{29}$$

> $\mathrm{sum11} := \mathrm{add}((\mathrm{Xi1}[i] - Xbar1)\cdot(\mathrm{Yi1}[i] - Ybar1), i = 1..N1)$:

> $\mathrm{sum21} := \mathrm{add}\left((\mathrm{Xbar1} - \mathrm{Xi1}[\mathrm{i}])^2, i = 1..N1\right)$:

> $b1 := \frac{sum11}{sum21}$:

> $Ycv1 := Vector([seq(Yi1[i] - b1\cdot(Xi1[i] - EX), i = 1..N1)])$:

> $100\cdot Mean(Ycv1)$

$$5.81905594736172693 40 \tag{30}$$

> $100\cdot StandardError(Mean, Ycv1)$;

$$0.07390481485700663209 5 \tag{31}$$

> ########## *optimal coefficient* b_{N2}, *price* **and** *standard* **error for** $N = 5.000$ **in** *Cents* #########

> $Xi2 := \mathrm{Xi}[1..5000]$:

> $Yi2 := Yi[1..5000]$:

> $Xbar2 := Mean(Xi2)$

$$Xbar2 := 0.04517591916345480671 0 \tag{32}$$

> $Ybar2 := Mean(Yi2)$

$$Ybar2 := 0.06510154866809093667 2 \tag{33}$$

> $\mathrm{sum12} := \mathrm{add}((\mathrm{Xi2}[i] - Xbar2)\cdot(\mathrm{Yi2}[i] - Ybar2), i = 1..N2)$:

> $\mathrm{sum22} := \mathrm{add}\left((\mathrm{Xbar2} - \mathrm{Xi2}[\mathrm{i}])^2, i = 1..N2\right)$:

> $b2 := \frac{sum12}{sum22}$:

> $Ycv2 := Vector([seq(Yi2[i] - b2\cdot(\mathrm{Xi2}[i] - EX), i = 1..N2)])$:

> $100\cdot Mean(Ycv2)$;

$$5.89955267987247225 42 \tag{34}$$

> $100\cdot StandardError(Mean, Ycv2)$;

$$0.14074026319443596162 \tag{35}$$

> ##

> ############################## *Calculating Variance Ratios* ############################

> ##

>

> $RatioArith := \frac{Variance(Cneutral)}{Variance(CArith)}$;

$$RatioArith := 133.97691625950460742 \tag{36}$$

> $RatioGeo := \frac{Variance(Cneutral)}{Variance(CGeo)}$;

$$RatioGeo := 131.21213138835458074 \tag{37}$$

> $RatioControlVariate := \frac{Variance(Cneutral)}{Variance(Ycv)}$;

$$RatioControlVariate := 40.115277827046249039 \tag{38}$$

>

```
> ####################################################################################
> ########################### arithmetic average Asian put option ########################
> ####################################################################################
> restart;
> with(Finance) :
> with(Statistics) :
>
> #################################### initialize values ####################################
> W := WienerProcess( ) :
> T := 1 :
> S := 50 :
> sigma := 0.25 :
> r := 0.05 :
> K := 35 :
> N := 100000 :
> N1 := 20000 :
> N2 := 5000 :
> n = 252 :
> ######################### N sample paths X of a Brownian motion #######################
> X := SamplePath(W(t), t = 0 ..1, timesteps = 252, replications = N) :
>
> ########################### geometric and arithmetic average #########################
```

$$> SGeo := Vector\left(\left[seq\left(S\cdot\left(mul\left(e^{\frac{\left(r-\frac{\sigma^2}{2}\right)\cdot(j-1)}{252}+\mathrm{sigma}\cdot X[i,j]}, j=1..253\right)\right)^{\frac{1}{253}}, i=1..N\right)\right]\right):$$

$$> SArith := Vector\left(\left[seq\left(\frac{S\cdot add\left(e^{\frac{\left(r-\frac{\sigma^2}{2}\right)\cdot(j-1)}{252}+\mathrm{sigma}\cdot X[i,j]}, j=1..253\right)}{253}, i=1..N\right)\right]\right):$$

```
>
> ####################################################################################
> ############################## usual Monte Carlo estimator #############################
> ####################################################################################
>
> ########### vector containing N option prices corresponding to N sample paths ############
```

$$> Cneutral := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max(\mathrm{K-SArith[i]}, 0), i=1..N\right)\right]\right):$$

```
>
> ################### price and standard error for N=100.000 in Cents ##################
> 100· Mean(Cneutral);
                                   0.4442306452                                   (1)
> 100· StandardError(Mean, Cneutral);
                                0.0313849988042821                                (2)
> ################### price and standard error for N=20.000 in Cents ###################
> C1neutral := Cneutral[1 ..20000] :
```

> $100 \cdot Mean(C1neutral);$

$$0.4751316700 \quad (3)$$

> $100 \cdot StandardError(Mean, C1neutral);$

$$0.0720564485276888 \quad (4)$$

> ##################### price and standard error for N=5.000 in Cents ####################

> $C2neutral := Cneutral[1..5000]:$

> $100 \cdot Mean(C2neutral);$

$$0.3463105204 \quad (5)$$

> $100 \cdot StandardError(Mean, C2neutral);$

$$0.106391575173950 \quad (6)$$

> ###

> ######### Monte Carlo estimator using Importance Sampling and the geometic drift #########

> ###

> $aGeo := \frac{\text{sigma}}{T}:$

> $cGeo := \frac{K}{S} \cdot e^{-\frac{\left(r - \frac{\sigma^2}{2}\right) \cdot T}{2}}:$

>

> #################################### compute beta ####################################

> $equ := aGeo \cdot b \cdot T^3 + 3 \cdot \ln\left(\frac{(b - aGeo)}{cGeo \cdot b}\right) = 0:$

> $betaGeo := fsolve(equ, b, b = -10000..0);$

$$betaGeo := -4.980337138 \quad (7)$$

> ######################### aymptotically optimal change of drift #########################

> $h := t \rightarrow -\frac{betaGeo}{2} \cdot t^2 + betaGeo \cdot T \cdot t:$

> $h1 := t \rightarrow betaGeo \cdot (T - t):$

>

> ########################### computation of the rescaling terms ##########################

> $Rescale1 := Vector\left(\left[seq\left(betaGeo \cdot \frac{add(X[i,j], j = 1..253)}{253}, i = 1..N\right)\right]\right):$

> $Rescale2 := e^{-0.5 \cdot \int_0^T (h1(t))^2 \, dt}:$

> ########### vector containing N option prices corresponding to N sample paths ###########

> $CGeo := Vector\left(\left[seq\left(e^{-r \cdot T} \cdot \max\left(K - \frac{add\left(S \cdot e^{\frac{\left(r - \frac{\sigma^2}{2}\right) \cdot (j-1)}{252} + \text{sigma} \cdot \left(h\left(\frac{(j-1)}{252}\right) + X[i,j]\right)}, j = 1..253\right)}{253}, 0\right) \cdot e^{-Rescale1[i]}\right.\right.\right.$

$\cdot Rescale2, i = 1..N \Big) \Big] \Big] \Big) :$

> ################### *price and standard error for N=100.000 in Cents* ###################

> $100 \cdot Mean(CGeo);$

$$0.5186146082 \tag{8}$$

> $100 \cdot StandardError(Mean, CGeo);$

$$0.00197185470970393 \tag{9}$$

> ################### *price and standard error for N=20.000 in Cents* ###################

> $C1Geo := CGeo[1..20000]:$

> $100 \cdot Mean(C1Geo);$

$$0.5190723783 \tag{10}$$

> $100 \cdot StandardError(Mean, C1Geo);$

$$0.00440892932215184 \tag{11}$$

> #################### *price and standard error for N=5.000 in Cents* ###################

> $C2Geo := CGeo[1..5000]:$

> $100 \cdot Mean(C2Geo);$

$$0.5210854301 \tag{12}$$

> $100 \cdot StandardError(Mean, C2Geo);$

$$0.00882799166940856 \tag{13}$$

>

> ##

> ######### *Monte Carlo estimator using Importance Sampling and the arithmetic drift* #########

> ##

> $a := \text{sigma}:$

> $b := r - \frac{\sigma^2}{2}:$

> $c := \frac{K \cdot T}{S}:$

> $d := \frac{S}{T}:$

> $\text{beta} := -0.0005;$

$$\beta := -0.0005 \tag{14}$$

> $\text{delta} := -1.00081;$

$$\delta := -1.00081 \tag{15}$$

>

> ######################### *aymptotically optimal change of drift* #########################

> $h := t \rightarrow \frac{\text{beta} - b}{a} \cdot t - \frac{2}{a} \cdot \ln\left(\frac{e^{\text{beta} \cdot t} + \text{delta}}{1 + \text{delta}}\right):$

> $h1 := t \rightarrow \frac{\text{beta} - b}{a} - \frac{2}{a} \cdot \frac{\text{beta} \cdot e^{\text{beta} \cdot t}}{e^{\text{beta} \cdot t} + \text{delta}}:$

>

> ########################## *computation of the rescaling terms* ##########################

$$> Rescale1 := Vector\left(\left[seq\left(h1(T)\cdot X[i,253] + T\cdot\frac{2\cdot\beta^2\cdot\delta}{a}\cdot\frac{add\left(\frac{e^{\frac{\text{beta}\cdot(j-1)}{252}}}{\left(e^{\frac{\text{beta}\cdot(j-1)}{252}}+\text{delta}\right)^2}X[i,j], j=1..253\right)}{253}, i=1..N\right)\right]\right):$$

$$> Rescale2 := e^{-0.5\cdot\int_0^T (h1(t))^2\,dt}:$$

> ########### *vector containing N option prices corresponding to N sample paths* ############

$$> CArith := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max\left(K - \frac{add\left(S\cdot e^{\frac{\left(r-\frac{\sigma^2}{2}\right)\cdot(j-1)}{252}+\text{sigma}\cdot\left(h\left(\frac{(j-1)}{252}\right)+X[i,j]\right)}, j=1..253\right)}{253}, 0\right)\cdot e^{-Rescale1[i]}\cdot Rescale2, i=1..N\right)\right]\right):$$

> #################### *price and standard error for N=100.000 in Cents* ###################

> 100· *Mean*(*CArith*);

0.5276059876 **(16)**

> 100· *StandardError*(*Mean*, *CArith*);

0.009813436644412574 **(17)**

> #################### *price and standard error for N=20.000 in Cents* ####################

> *C1Arith* := *CArith*[1..20000] :

> 100· *Mean*(*C1Arith*);

0.5126037718 **(18)**

> 100· *StandardError*(*Mean*, *C1Arith*);

0.0168597282865647 **(19)**

> ##################### *price and standard error for N=5.000 in Cents* ####################

> *C2Arith* := *CArith*[1..5000] :

> 100· *Mean*(*C2Arith*);

0.4942643426 **(20)**

> 100· *StandardError*(*Mean*, *C2Arith*);

0.0313799322651236 **(21)**

>

> ##

> ############### *Monte Carlo estimator using the method of control variates* ##############

> ##

> $m := 252 + 1:$

> $z := \text{sigma}\cdot\sqrt{\frac{2\cdot m+1}{6\cdot(m+1)}}:$

> $\text{rho} := \frac{\left(r - \frac{\sigma^2}{2} + z^2\right)}{2}:$

> $\text{phi} := RandomVariable(Normal(0, 1)):$

>

> $d1 := \frac{\ln\left(\frac{S}{K}\right) + \left(\text{rho} + \frac{z^2}{2}\right)\cdot T}{\sqrt{T}\cdot z}:$

> $d2 := \frac{\ln\left(\frac{S}{K}\right) + \left(\text{rho} - \frac{z^2}{2}\right)\cdot T}{\sqrt{T}\cdot z}:$

>

> ############# *computing Xi , Yi , Xbar, Ybar, EX as described in the method* ##############

> $\text{Xi} := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max(\text{K}-\text{SGeo}[\text{i}], 0), i = 1..N\right)\right]\right):$

> $Yi := Vector\left(\left[seq\left(e^{-r\cdot T}\cdot\max(\text{K}-\text{SArith}[\text{i}], 0), i = 1..N\right)\right]\right):$

> $Xbar := Mean(\text{Xi})$

$$Xbar := 0.007260217225 \quad \textbf{(22)}$$

> $Ybar := Mean(Yi)$

$$Ybar := 0.004442306452 \quad \textbf{(23)}$$

> $EX := e^{(\text{rho}-r)\cdot T}\cdot\left(K\cdot e^{-\text{rho}\cdot T}\cdot CDF(\text{phi}, -d2) - S\cdot CDF(\text{phi}, -d1)\right):$

> $simplify(EX)$

$$0.008181550000 \quad \textbf{(24)}$$

> ######### *optimal coefficient* b_N *, price* **and** *standard* **error for** $N = 100.000$ **in** *Cents* ########

> $sum1 := \text{add}((\text{Xi}[i] - Xbar)\cdot(\text{Yi}[i] - Ybar), i = 1..N):$

> $\text{sum2} := \text{add}\left((\text{Xbar} - \text{Xi}[\text{i}])^2, i = 1..N\right):$

> $b := \frac{sum1}{sum2};$

$$b := 0.717657778627460 \quad \textbf{(25)}$$

> $Ycv := Vector([seq(Yi[i] - b\cdot(\text{Xi}[i] - EX), i = 1..N)]):$

> $100\cdot Mean(Ycv)$

$$0.5103506326 \quad \textbf{(26)}$$

> $100\cdot StandardError(Mean, Ycv);$

$$0.00749325129650342 \quad \textbf{(27)}$$

> ######### *optimal coefficient* b_{N1} *, price* **and** *standard* **error for** $N = 20.000$ **in** *Cents* #########

> $Xi1 := \mathrm{Xi}[1..20000]:$

> $Yi1 := Yi[1..20000]:$

> $Xbar1 := Mean(Xi1);$

$$Xbar1 := 0.007754967450 \quad (28)$$

> $Ybar1 := Mean(Yi1);$

$$Ybar1 := 0.004751316700 \quad (29)$$

> $\mathrm{sum11} := \mathrm{add}((\mathrm{Xi1}[i] - Xbar1) \cdot (\mathrm{Yi1}[i] - Ybar1), i = 1..N1):$

> $\mathrm{sum21} := \mathrm{add}\big((\mathrm{Xbar1} - \mathrm{Xi1}[\mathrm{i}])^2, i = 1..N1\big):$

> $b1 := \dfrac{sum11}{sum21}:$

> $Ycv1 := Vector([seq(Yi1[i] - b1 \cdot (Xi1[i] - EX), i = 1..N1)]):$

> $100 \cdot Mean(Ycv1)$

$$0.5052898665 \quad (30)$$

> $100 \cdot StandardError(Mean, Ycv1);$

$$0.0176975823201041 \quad (31)$$

> ########## *optimal coefficient* b_{N2} , *price* **and** *standard* **error for** $N = 5.000$ **in** *Cents* #########

> $Xi2 := \mathrm{Xi}[1..5000]:$

> $Yi2 := Yi[1..5000]:$

> $Xbar2 := Mean(Xi2)$

$$Xbar2 := 0.006921757313 \quad (32)$$

> $Ybar2 := Mean(Yi2)$

$$Ybar2 := 0.003463105204 \quad (33)$$

> $\mathrm{sum12} := \mathrm{add}((\mathrm{Xi2}[i] - Xbar2) \cdot (\mathrm{Yi2}[i] - Ybar2), i = 1..N2):$

> $\mathrm{sum22} := \mathrm{add}\big((\mathrm{Xbar2} - \mathrm{Xi2}[\mathrm{i}])^2, i = 1..N2\big):$

> $b2 := \dfrac{sum12}{sum22}:$

> $Ycv2 := Vector([seq(Yi2[i] - b2 \cdot (\mathrm{Xi2}[i] - EX), i = 1..N2)]):$

> $100 \cdot Mean(Ycv2);$

$$0.4219368849 \quad (34)$$

> $100 \cdot StandardError(Mean, Ycv2);$

$$0.0355755211768194 \quad (35)$$

> ##

> ############################## *Calculating Variance Ratios* ###########################

> ##

>

> $RatioArith := \dfrac{Variance(Cneutral)}{Variance(CArith)};$

$$RatioArith := 10.2282653632800 \quad (36)$$

> $RatioGeo := \dfrac{Variance(Cneutral)}{Variance(CGeo)};$

$$RatioGeo := 253.334541492767 \quad (37)$$

> $RatioControlVariate := \dfrac{Variance(Cneutral)}{Variance(Ycv)};$

(38)

$$RatioControlVariate := 17.5429908908226 \tag{38}$$

>

>

>

Bibliography

[1] den Hollander, F.: Large Deviations, Fields Institute Monographs, 2000

[2] Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications, Springer-Verlag, 2nd Edition, 1998

[3] Glasserman, P.: Monte Carlo Methods in Financial Engineering, Springer, 2004

[4] Guasoni, P., Robertson, S.: Optimal Importance Sampling with explicit formulas in continuous time, Springer, 2007

[5] Glasserman, P., Heidelberger, P., Shahabuddin, P.: Asymptotically optimal importance sampling and stratification for pricing path dependent options , 1999

[6] Gelfand, M., Fomin, S. V.: Calculus of Variations, Prentice-Hall, 1963

[7] Musiela, M., Rutkowski, M.: Martingale Methods in Financial Modeling, Springer-Verlag, 1997.

[8] Øksendal, B.: Stochastic Differential Equations, Springer-Verlag, 6th edition, 2003.

[9] Shreve, S. E.: Stochastic Calculus for Finance II, Springer-Verlag, 2004.

[10] Hull, J. C.: Options, Futures and other Derivatives, Pearson Prentice Hall, 8th edition, 2011.

[11] Zhang, H.: Princing Asian Options using Monte Carlo Methods, Uppsala Universitet, 2009.