universität wien

DISSERTATION

Titel der Dissertation

„Efficient calculation of the Greeks:

An application of Measure Valued Differentiation“

Verfasser

Philipp Thoma

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, im Jänner 2012

Studienkennzahl lt. Studienblatt:	A 094 136
Dissertationsgebiet lt. Studienblatt:	Statistik und Operations Research
Betreuerin / Betreuer:	O.Univ.-Prof. Mag. Dr. Georg Ch. Pflug

Contents

1 Preface **7**

2 Derivative contracts and underlying stochastic processes **9**
2.1 Options . . . 9
2.2 Greeks . . . 10
2.3 Stochastic differential equations . . . 11
2.4 Processes with stationary independent increments (Lévy processes) . . . 13
2.4.1 Basics . . . 13
2.4.2 Brownian Motion with drift μ and diffusion σ . . . 15
2.4.3 Poisson process . . . 16
2.4.4 Compound Poisson process . . . 18
2.4.5 Gamma process . . . 20
2.4.6 Variance Gamma process . . . 22
2.4.7 Construction of Lévy processes . . . 24

3 Lévy processes in finance **27**
3.1 Market model . . . 27
3.2 Change of measure . . . 28
3.3 Exponential Lévy processes . . . 32
3.3.1 Geometric Brownian Motion model . . . 32
3.3.2 Variance Gamma model . . . 34
3.3.3 Compound Poisson model . . . 34

4 Sensitivity (Gradient) estimation methods **37**
4.1 Basic Problem . . . 37
4.2 Finite difference (FD) method . . . 38
4.3 Pathwise method . . . 39
4.4 Score-Function method . . . 43
4.5 Malliavin Calculus . . . 44
4.5.1 Introduction . . . 44
4.5.2 Malliavin Derivative . . . 44
4.5.3 The Skorohod Integral . . . 45
4.5.4 Integration by parts procedure . . . 46
4.6 Measure valued differentiation . . . 47
4.6.1 Introduction . . . 47
4.6.2 Examples for weak derivatives . . . 49
4.6.3 MVD via the characteristic function . . . 51

4.6.4 The function set $\mathcal{H}$ 54
4.6.5 The Measure Valued Derivative for Markov processes and its estimation 57
4.6.6 Algorithms MVD 59
4.6.6.1 The exact estimator (MVDe) 59
4.6.6.2 The randomized estimator (MVDr) 59
4.6.6.3 The compromise estimate (MVDc(k)) 61
4.6.7 Variance reduction via coupling 62
4.6.8 MVD for path dependent payoff functions 63
4.6.8.1 Estimation via joint transition densities 63
4.6.8.2 Estimation via Operators 69

5 Implementation of MVD for exponential Lévy processes **73**
5.1 Geometric Brownian motion model 73
5.1.1 Rho (ρ): Sensitivity w. r. t. r 73
5.1.2 Vega (ν): Sensitivities w.r.t. σ 78
5.1.3 Delta (Δ): Sensitivities w.r.t. $S(0)$ 80
5.2 Poisson model 82
5.3 Gamma model 83
5.4 The Variance Gamma model 84
5.5 Compound Poisson model 85
5.5.1 Sensitivity w.r.t. b_1 85
5.5.2 Sensitivity w.r.t. λ 86

6 Appendices **89**

Bibliography **97**

1 Preface

In Finance, simulation methods have become standard methods for pricing various, but especially exotic financial contracts under various model assumptions. This work focuses on the estimation of sensitivities of option prices w.r.t. changes of distributional parameters, the so called "Greeks", when the underlying stochastic model follows some Markov process of exponential Lévy-type. Price sensitivities are used in measuring and managing risk and find also application in hedging-strategies.
Lévy processes are widely used and analyzed as models of asset prices, interest rates, exchange rates, commodities and other financial variables. In the classical Black-Scholes framework, there exist closed formulas for the Greeks of plain Vanilla Call or Put Options. But if we turn to more complicated payoff functions (e.g. Exotic Options) or if we consider other driving Lévy processes than Brownian motion, numerical approaches relying on Monte Carlo methods are needed to estimate both, the values and their sensitivities.

Suppose that $S_\theta(t)$ describes the random price process of the underlying and H is the payoff function of an option on this underlying with maturity T. θ is some parameter governing the distribution of S. The fair option price is

$$\mathbb{E}_Q[e^{-rT}H(S_\theta(T))],$$

where Q is a probability law which makes the discounted process

$$\tilde{S}_\theta(t) = e^{-rt}S_\theta(t) \tag{1.0.1}$$

a martingale. Our goal is to calculate

$$\frac{\partial}{\partial\theta}\mathbb{E}_Q[e^{-rT}H(S_\theta(T))]. \tag{1.0.2}$$

Since closed form expressions do not exist (except for very special cases), estimations of (1.0.2) with high accuracy are needed.

There exist several Monte Carlo techniques for estimating the Greeks. Each of those methods have their individual advantages and disadvantages, depending on the model of the stock price and the structure of the payoff function.

The most simple is the *finite difference (FD) approximation*: The price is estimated under $\theta + h$ and under θ and the difference of the estimates is divided by h. Even if the variance of the FD estimate is reduced by taking highly correlated estimates, there is still the bias issue: FD estimates are biased. The more sophisticated methods are subdivided into methods which assume parameterized integrands

$$\theta \mapsto \int H(\theta,\omega)d\nu(\omega)$$

and methods with parametrized integrators

$$\theta \mapsto \int H(\omega) d\nu_\theta(\omega),$$

see [33]. The *pathwise method (Infinitesimal Pertubation Analysis IPA)* is a method of the first sort, see [12]. Other direct methods belong to the second group, as the *Likelihood ratio method* or *Score Function method*, [12], the *Malliavin calculus*, [37, 26, 39], and the *Measure valued differentiation*, [3, 4, 5, 18, 41].
The main contribution of this work is to demonstrate how the MVD method may by applied for a variety of different stock-price models and different, notably nondifferentiable payoff functions.

The work is organized as follows. In section 2 we give a general introduction to options, their sensitivities (*Greeks*) and stochastic processes in financial context. Section 3 introduces Lévy processes which we use to model the evolution of stock prices. According to that, different Lévy-type models are presented, beside the well known Geometric Brownian Motion model. In Section 4 we discuss different types of simulation methods to estimate the *Greeks,* first and foremost we introduce the measure valued differentiation method and its extension for Markov processes. This section also demonstrates how the method can be applied to the sensitivity estimation for path-dependent payoffs, such as Lookback and Asian Options. The last section is dedicated to the implementation of MVD for exponential Lévy processes and comprised some numerical examples about the performance of the estimation method for path-independent and path-dependent payoffs.

At that point i want to thank Prof. Pflug who initiated me to this subject and was always there for fundamental discussions and helpful suggestions. During my period as a PhD student i got to know him not only as an ambitious advisor, but also as a person who knows how to create a pleasant working atmosphere among our little research group, which was invigorated during a multitude of activities outside the institute. I also want to thank Prof. Heidergott, who generously gave valuable remarks on an earlier version of this work.

Philipp Thoma

Vienna, 18.August 2011

2 Derivative contracts and underlying stochastic processes

This section provides the fundamental background of this work. It is dedicated to derivative contracts and their sensitivities, as well as to stochastic differential equations which play an important role in mathematical finance. Further we discuss Lévy processes, which build the basis for modelling the underlying.
Derivative contracts are instruments that derive their values from other underlying variables, e.g. stocks [1] and shares. The most common derivatives are futures, options and swaps. This work focuses on options, especially on the determination of sensitivities of option prices with respect to small perturbations of the dynamics of the underlying asset. Those sensitivities are called "*Greeks*" in finance world. For further information of the mechanics of option markets see [7].

2.1 Options

While *underlying stocks* as shares, bonds, currencies or commodities give the holder the direct posession of a good, an *option* gives the holder the right, but not the obligation, to buy or to sell a specified amount of a stock (*asset)* on or before the options expiration time (*maturity*) for a price specified in advance, the *strike price.*
Options that give the holder the right to buy the underlying security at the strike price are called *call options* whereas *put options* give the holder the right to sell. If the buyer of an option choose to exercise his right to buy/sell, the seller is obliged to sell/buy the asset at the agreed price.
There exist many different types of options. For example an *European Option* gives the holder the right to buy/sell the underlying asset only at maturity while an *American Option* allows this at any trading day on or before maturity. But this are just the standard option styles. Beside of this, there exist options with more complicated payoff functions, the socalled "*Exotic" Options.* Exotic Options have a payoff function which is different from the usual Call or Put Option, e.g. Digital Options or Lookback Options. Special Exotic Options are one of the *rainbow type.* Rainbow Options are linked to several underlyings.
Here are some examples of Rainbow Options, see [22], where $S_i(T)$ describes the price of stock i at maturity time T.

[1] A stock (also known as a share or as an equity) represents ownership interests in a company.

Type	Payoff
Margrabe Option	$max(S_1(T) - S_2(T), 0)$
Better-off Option	$max(S_1(T), ..., S_n(T))$
Worse-off Option	$min(S_1(T), ..., S_n(T))$
Maximum Option	$max(max(S_1(T), ..., S_n(T)) - K, 0)$
Minimum Option	$max(min(S_1(T), ..., S_n(T)) - K, 0)$
Spread Option	$max(S_2(T) - S_1(T) - K, 0)$
Basket average Option	$max\left(\frac{S_1(T)+...+S_n(T)}{n} - K, 0\right)$
Multi-strike Option	$max(S_1(T) - K_1, ..., S_n(T) - K_n, 0)$
Pyramid rainbow Option	$max(\|S_1(T) - K_1\| + ... + \|S_n(T) - K_n\| - K)$
Madonna rainbow Option	$max\sqrt{(S_1(T) - K_1)^2 + ... + (S_n(T) - K_n)^2 - K, 0}$

Table 2.1: List of Rainbow Options.

Options can be used as insurance against price hikes, or for speculations on increasing or decreasing stock prices. Hedgers for example, use put options to limit potential loss if the underlying stock is assumed to decline. Speculators hope to anticipate the price movement of the underlying to make large profits.
The pricing of options plays a very important role in the financial sector. Let $g(S(T)) = C(T)$ be the options cashflow. The fair price of an option today $(\Pi_0(C_T))$, is the minimal initial capital needed to replicate the cashflow of the option by trading with the riskless bond and the underlying. That is equivalent to the assumption of *No Arbitrage* or the existence of a risk neutral measure, which will be explained later in section 3.1.

2.2 Greeks

A specifically important problem in finance is the estimation of sensitivities of option prices, the socalled *"Greeks"*. These sensitivities are used in measuring and managing risk and find also application in hedging-strategies. In a complete market, any payoff can be synthesized through a trading strategy. The risk in a short call position in an option, for example, is offset by a *delta-hedging* strategy of holding delta units of each underlying asset. Thereby is delta the partial derivative of the options price with respect to the current price of that underlying asset, see [7] for further hedging strategies.
Due to the fact that those sensitivites cannot be observed directly in the market, they need to be calculated. In the classical Black-Scholes framework, where the stock price follows a log-normal distribution, there exist closed formulas for the Greeks of plain Vanilla Call or Put Options. But if we consider other driving processes or more complicated payoff functions like the Rainbow Options, numerical approaches relying on Monte Carlo methods are needed to estimate both, the values and their the sensitivities. There exist several simulation methods based on Monte Carlo techniques for estimating those sensitivities, which will be discussed in section 4.

Type of Greeks

Some classical sensitivites are

- Δ: The (first) derivative of the options price w.r.t. the price of the underlying
- Γ: The second derivative of the options price w.r.t. the price of the underlying
- Θ: The derivative of the options price w.r.t. the time of maturity
- ν: The derivative of the options price w.r.t. the volatility
- ρ: The derivative of the options price w.r.t. the risk-free interest rate

see [7]. However this list is not complete, because it orients itself on the Geometric Brownian motion model with non-stochastic risk free interest rate (which is only in a large class of models). Additional parameters will appear e.g. in stochastic interest rate models or in Lévy-type models.

2.3 Stochastic differential equations

Stochastic differential equations are used to describe the dynamics of random phenomena in a wide range of natural sciences, as so they do in finance.
A stochastic differential equation (SDE) is a differential equation in which one ore more terms follow a stochastic process. The solution of such a SDE is therefore also a stochastic process. In mathematical finance such an SDE is often of the form

$$dX(t) = f(X(t), t)dt + \sigma(X(t), t)dW(t) \tag{2.3.1}$$

which is short for

$$X(t) - X(0) = \int_0^t f(X(s), s)ds + \int_0^t \sigma(X(s), s)dW(s)$$

where $W(s)$ denotes a Wiener Process[2]. Under certain regularity conditions, see [31], there exists a solution of 2.3.1.
In the homogeneous, autonomous case (f, σ independent of time), the SDE is described by

$$dX(t) = f(X(t))dt + \sigma(X(t))dW(t)$$

Famous Examples for SDE's in mathematical finance are

- Vasicek model, Ornstein-Uhlenbeck model: $dX(t) = a(\mu - X(t))dt + \sigma dW(t)$

[2] A Wiener process is also often called a standard Brownian motion described by $B(s)$. For the definition of a Wiener process see [12] page 79.

- Black-Scholes-model[3]: $dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$
- Cox-Ingersoll-Ross model: $dX(t) = a(\mu - X(t))dt + \sqrt{X(t)}\sigma dW(t)$
- Mean reverting geometric Brownian motion: $dX(t) = a(\mu - X(t))dt + \sigma X(t)dW(t)$

Sensitivities Let uns consider a homogeneous, autonomous SDE. Let us further assume we are interested in the sensitivity of the SDE with respect to the starting point. In addition $X(0) = x$ and f, σ are smooth and continuous functions. Then we can approximate by

$$X(h) \approx x + f(x)h + \sigma(x)Z(h)$$

where $Z(h) \sim N(0, h)$. Therefore we get for time $t + h$

$$X(t+h) \approx X(t) + f(X(t))h + \sigma(X(t))Z(h).$$

Let $D_x(t) = \frac{\partial X(t)}{\partial x}$ and $D_x(0) = 1$. Then we get for the derivative w.r.t x

$$D_x(t+h) \approx D_x(t) + f'(X(t))\frac{\partial X(t)}{\partial x}h + \sigma'(X(t))\frac{\partial X(t)}{\partial x}Z(h).$$

For the SDE's this yields

$$\begin{aligned} dX(t) &= f(X(t))dt + \sigma(X(t))dW(t) \\ dD_x(t) &= f'(X(t))D_x(t)dt + \sigma'(X(t))D_x(t)dW(t) \end{aligned}$$

To get the derivative w.r.t. x we first simulate $X(t)$, which is then used in the simulation of $D_x(t)$.

If the SDE is additionally dependent on some other parameter ϑ

$$dX_\vartheta(t) = f_\vartheta(X_\vartheta(t))dt + \sigma_\vartheta(X_\vartheta(t))dW(t)$$

we can of course also look for the sensitivity of the SDE w.r.t. ϑ. Let

$$\begin{aligned} f'_\vartheta(x) &= \frac{\partial}{\partial x} f_\vartheta \\ \dot{f}_\vartheta(x) &= \frac{\partial}{\partial \vartheta} f_\vartheta \\ \frac{\partial}{\partial \vartheta} X_\vartheta(t) &= D_\vartheta(t) \end{aligned}$$

[3] The model was first established by Fischer Black and Myron Scholes. Robert C. Merton was also involved in the elaboration. Merton and Scholes received 1997 the Nobel Prize in Economics. Cause of the death of Black in 1995, he was mentioned as a contributor by the Swedish academy

and

$$X_\vartheta(t+h) \approx X_\vartheta(t) + f_\vartheta(X_\vartheta(t))h + \sigma_\vartheta(X_\vartheta(t))Z(h).$$

Then

$$D_\vartheta(t+h) = D_\vartheta(t) + \left[\dot{f}_\vartheta(X_\vartheta(t)) + f'_\vartheta(X_\vartheta(t))D_\vartheta(t)\right] h + \left[\dot{\sigma}_\vartheta(X_\vartheta(t)) + \sigma'_\vartheta(X_\vartheta(t))D_\vartheta(t)\right] Z(h)$$

and for the differentiated SDE we get

$$dD_\vartheta(t) = \left[\dot{f}_\vartheta(X_\vartheta(t)) + f'_\vartheta(X_\vartheta(t))D_\vartheta(t)\right] dt + \left[\dot{\sigma}_\vartheta(X_\vartheta(t)) + \sigma'_\vartheta(X_\vartheta(t))D_\vartheta(t)\right] dW(t) \tag{2.3.2}$$

Example (Geometric Brownian motion).

$$dX(t) = \underbrace{\left(\mu - \frac{1}{2}\sigma^2\right) X(t)}_{f_\vartheta(X(t))} dt + \sigma X(t) dW(t)$$

where the differentiated GBM with $D_\mu(0) = 0$ is

$$dD_\mu(t) = \left[X_\mu(t) + \left(\mu - \frac{1}{2}\sigma^2\right) D_\mu(t)\right] dt + \sigma D_\mu(t) dW(t)$$

with solution $D_\mu(t) = t \cdot X(t)$.

2.4 Processes with stationary independent increments (Lévy processes)

In Finance, the evolution of stock prices are usually modelled by processes with stationary independent increments, socalled *Lévy processes*. Before we observe these processes we have to do some relevant basics which can be found in more detail in [32, 38, 35].

2.4.1 Basics

Let $X(t) : \Omega \to \mathbb{R}$ be a stochastic Process.

Definition. A stochastic process has independent increments if

$$X(t+h) - X(t) \qquad \text{independent of} \quad \mathcal{F}_t$$

with the sigma-algebra $\mathcal{F}_t = \sigma(X_s : 0 \leq s \leq t)$.

Definition. A stochastic process has stationary increments if

$$X(t+h) - X(t) \overset{\mathcal{L}}{=} X(s+h) - X(s) \qquad \forall t, s$$

where $\overset{\mathcal{L}}{=}$ means equal in distribution.

Examples for such processes are

- Wiener process:

$$W(0) = 0$$
$$W(t+h) - W(t) \sim N(0,h)$$

- Poisson process:

$$N(0) = 0$$
$$N(t+h) - N(t) \sim \text{Poisson}(\lambda h)$$

- Gamma process:

$$G(0) = 0$$
$$G(t+h) - G(t) \sim \Gamma(ah,b)$$

All Lévy processes are Markovian, in addition they are martingales if

$$\mathbb{E}\left(X(t+h) - X(t)\right) = 0.$$

A Lévy process has independent and stationary increments. It is fully characterized by the characteristic function φ of its (infinitely divisible[4]) increment distribution

$$\mathbb{E}\left[exp\left(iu\left(X(t+1) - X(t)\right)\right)\right] = \varphi(u).$$

By the property of infinite divisibility, the characteristic function of the increment in time interval Δ is just the Δ-th power of ϕ

$$\mathbb{E}\left[exp\left(iu\left(X(t+\Delta) - X(t)\right)\right)\right] = \left(\varphi(u)\right)^{\Delta}.$$

With the following theorem we can characterize Lévy processes via their characteristic function. The theorem can be found in [30, 38].

Theorem 1 (Lévy-Chinchin representation). *$\varphi_X(u)$ is the characteristic function of a random variable X with infinitely divisible distribution iff it is of the form*

$$\varphi_X(u) = exp\left(aiu - \frac{1}{2}\sigma^2 u^2 + \int\limits_{\mathbb{R}\backslash\{0\}} \left(e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}}\right) d\nu(x)\right) \tag{2.4.1}$$

with the Lévy triplet $(a, \sigma^2, d\nu(x))$ with $a \in \mathbb{R}, \sigma^2 \geq 0$ .Usually a is denoted as the shift, σ^2 as the diffusion and ν as the socalled jump measure or Lévy measure. *The representation by 2.4.1 is unique and conversely, for any choice of a, σ^2 and $d\nu(x)$ there exists an infintely divisible distribution μ having characteristic function 2.4.1.*

[4]see 6 in Appendix

The cumulant characteristic function $\psi_X(u) = \log \varphi_X(u)$ is sometimes called the characteristic exponent of X.

Definition. $X(t)$ is a Lévy process, iff

- $X(t)$ is a process with stationary, independent increments or
- $X(t)$ is a process with infinitely divisible distribution where $X(t+\Delta) - X(t)$ has a distribution with characteristic function $[\varphi(u)]^\Delta$

2.4.2 Brownian Motion with drift μ and diffusion σ

Definition. A process $X(t)$ is called a Brownian motion with drift μ and diffusion σ if $\frac{X(t)-\mu t}{\sigma}$ is a Brownian motion with constant $\mu, \sigma > 0$. Then it is of the form

$$X(t) = \mu t + \sigma W(t)$$

and we write $X(t) \sim BM(\mu t, \sigma^2 t)$.

$X(t)$ can also be defined through an SDE of the form

$$dX(t) = \mu dt + \sigma dW(t)$$

and in the not autonomous case

$$dX(t) = \mu(t)dt + \sigma(t)dW(t).$$

The increment distribution is:

$$X(t+1) - X(t) \sim N(\mu, \sigma^2)$$

$N(\mu, \sigma^2)$ has density

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

and characteristic function

$$\varphi(u) = exp\left[iu\mu - \frac{\sigma^2}{2}u^2\right].$$

Distribution of increment in time interval Δ:

$$X(t+\Delta) - X(t) \sim N(\Delta\mu, \Delta\sigma^2)$$

with characteristic function

$$[\varphi(u)]^\Delta = exp\left[iu\Delta\mu - \frac{\Delta\sigma^2}{2}u^2\right].$$

Lévy triplet:

$$[a = \mu, \sigma^2, 0]$$

Additional facts: The paths of the Brownian motion are nowhere differentiable and of infinite variation within each interval. The paths are Hoelder continuous[5] with exponent $\beta \leq 1/2$. All moments of $X(t)$ exists. The Brownian motion is the only continuous Lévy process, i.e. it is the only process without the jump component $d\nu(x)$ in the characteristic function.

Moments:

$$\begin{aligned}
\mathbb{E}\left[(X(t+\Delta) - X(t))^1\right] &= i^{-1}\varphi'(0) = \frac{1}{i}\left(i\Delta\mu - \Delta\sigma^2 u) exp\left(iu\Delta\mu - \frac{\sigma^2}{2}\Delta u^2\right)\right)|_{u=0} \\
&= \Delta\mu \\
\mathbb{E}\left[(X(t+\Delta) - X(t))^2\right] &= i^{-2}\varphi''(0) = \Delta\sigma^2 + \Delta^2\mu^2
\end{aligned}$$

Simulation of Brownian Motion with drift μ and diffusion σ:
For the construction of a Brownian motion with drift and diffusion, we use the classical Euler scheme approximation (see [12]). For a fixed set of points $0 = t_0 < t_1 < ... < t_n$

$$X(t_{i+1}) = X(t_i) + \mu(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i} Z_{i+1} \text{ for } i = 0, ..., n-1$$

with $Z_1 ... Z_n \sim N(0,1)$ independent.
If we choose equidistant points for simulation we get

$$X(t_{i+1}) = X(t_i) + \mu\frac{1}{n} + \sigma\sqrt{\frac{1}{n}} Z_{i+1} \text{ for } i = 0, ..., n-1.$$

Figure 2.4.1 shows 5 different paths of a Brownian motion, where $X(1) - X(0) \sim \text{BM}(0.1, 0.5)$, constructed with Euler scheme approximation simulated at a set of 100 points with $X(0) = 100$.

2.4.3 Poisson process

Increment distribution:

$$X(t+1) - X(t) \sim \text{Poisson}(\lambda)$$

with distribution

$$P\{\text{Poisson}(\lambda) = k\} = exp(-\lambda)\frac{\lambda^k}{k!}$$

[5] A function $f : U \to \mathbb{R}$ is Hoelder continuous wirh exponent β iff there exists a positive real C such that for all $x, y \in U : |f(x) - f(y)| \leq C|x-y|^\beta$.

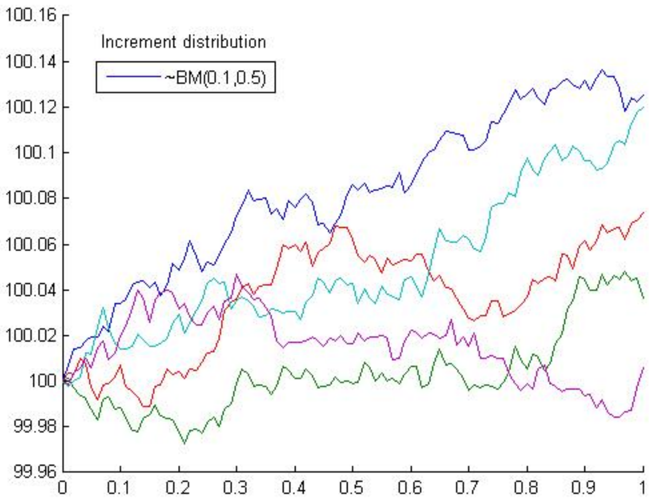


Figure 2.4.1: Brownian Motion with drift $\mu = 0.1$ and diffusion $\sigma = 0.5$.

and characteristic function

$$\varphi(u) = exp(\lambda(e^{iu} - 1)).$$

Distribution of increment in time interval Δ:

$$X(t + \Delta) - X(t) \sim \text{Poisson}(\Delta\lambda)$$

with characteristic function

$$[\varphi(u)]^{\Delta} = exp(\Delta\lambda(e^{iu} - 1))$$

Lévy triplet:

$$[0, 0, \lambda \cdot \delta_1]$$

where δ_1 is the point mass at 1.

Additional facts: The paths of the Poisson process are nondecreasing and take only integer values. All moments of $X(t)$ exist.

Moments:

$$\begin{aligned}
\mathbb{E}\left[(X(t+\Delta) - X(t))^1\right] &= i^{-1}\varphi'(0) = i^{-1}(i\Delta\lambda e^{iu} exp(\lambda\Delta(e^{iu} - 1))|_{u=0} \\
&= \Delta\lambda \\
\mathbb{E}\left[(X(t+\Delta) - X(t))^2\right] &= i^{-2}\varphi''(0) = \lambda\Delta + \lambda^2\Delta^2
\end{aligned}$$

Simulation of a Poisson Process: Euler approximation yields

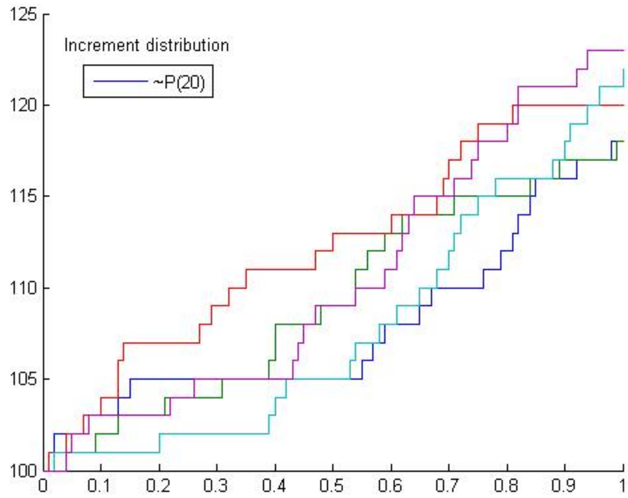


Figure 2.4.2: Poisson process with $\lambda = 20$.

$$X(t_{i+1}) = X(t_i) + P_{i+1} \text{ for } i = 0, ..., n-1$$

with independent $P_1, ..., P_n \sim \text{Poisson}(\lambda(t_{i+1}-t_i))$ respectively $P_1, ..., P_n \sim \text{Poisson}\left(\lambda\frac{1}{n}\right)$ if we simulate at equidistant points.

Figure 2.4.2 shows 5 paths of a Poisson process where $X(1) - X(0) \sim \text{Poisson}(20)$, constructed with Euler scheme approximation simulated at a set of 100 points and $X(0) = 100$.

2.4.4 Compound Poisson process

The Compound Poisson process is defined as

$$X(t) = \sum_{k}^{N(t)} Z_k \;\; t \geq 0.$$

where $N(t)$ is a Poisson process with intensity λ and (Z_k) are i.i.d. jump sizes, which are independent from $N(t)$. The increment distribution is

$$X(t+1) - X(t) \sim \sum_{k}^{N(1)} Z_k$$

with characteristic function

$$\begin{aligned}\varphi(u) &= \mathbb{E}\left(exp(isX(t))\right) = \mathbb{E}\left(exp\left(iu\sum_{j=1}^{N(t)} Z_j\right)\right)\\ &= \sum_{k=0}^{\infty}\mathbb{E}\left(exp\left(iu\sum_{j=1}^{k} Z_j\right)\right)\cdot P\left\{N_t = k\right\}\\ &= \sum_{k=0}^{\infty}\left[\varphi_Z(u)\right]^k e^{-\lambda t}\frac{(\lambda t)^k}{k!}\\ &= exp\left(\lambda t\left(\varphi_Z(u)-1\right)\right)\end{aligned}$$

where φ_Z is the characteristic function of the jump distribution.

Distribtution of increment with time interval Δ:

$$X(t+\Delta) - X(t) \sim \sum_{k=1}^{N(\Delta)} Z_k$$

where $N(\Delta)$ is Poisson$(\Delta\lambda)$ distributed.

We consider in particular jump distributions which are mixtures of Gamma distributions on the positive resp. negative half line, i.e.

$$Z_k = \begin{cases} V_k^+ & \text{with probability } p \\ V_k^- & \text{with probability } 1-p \end{cases}$$

with

$$\begin{aligned} V_k^+ &\sim \text{Gamma}(a_1, b_1)\\ V_k^- &\sim \text{Gamma}(a_2, b_2). \end{aligned}$$

independent of each other. This jump distribution has characteristic function $\varphi_Z(u) = p(1-iub_1)^{-a_1} + (1-p)(1+iub_2)^{-a_2}$. For this process we write

$$X(\cdot) \sim \text{CP}(\lambda, a_1, b_1, a_2, b_2, p).$$

Lévy triplet:

$$\begin{aligned}\exp\left(\lambda t\left(\varphi_Z(u)-1\right)\right) &= \exp\left(\lambda t\int_{-\infty}^{\infty}\left(\exp(iux)-1\right)d\mu(x)\right)\\ &= \exp\left(t\int_{-\infty}^{\infty}\left(\exp(iux)-1-iux\mathbf{1}_{\{|x|<1\}}+iux\mathbf{1}_{\{|x|<1\}}\right)d\nu(x)\right)\\ &= \exp\left(tiu\int_{-\infty}^{\infty}x\mathbf{1}_{\{|x|<1\}}d\nu(x) + t\int_{-\infty}^{\infty}\left(\exp(iux)-1-iux\mathbf{1}_{\{|x|<1\}}\right)d\nu(x)\right).\end{aligned}$$

where we consider $\nu(A) = \lambda \underbrace{P\{Z_1 \in A\}}_{\mu(A)}$.

Therefore we get for the Lévy triplet:

$$[\int_{-1}^{1} x d\nu(x), 0, d\nu(x)]$$

with $\mu(A) = P\{Z_1 \in A\}$.

Additonal facts: The paths of the compound Poisson process are of bounded variation . The m-th moment of X_t exists, iff the m-th moment of Z_1 exists.

Moments:

$$\begin{aligned}
\mathbb{E}\left[(X(t+\Delta) - X(t))^1\right] &= \frac{1}{i}\varphi'(0) = \frac{1}{i}\lambda\Delta\varphi_Z'(0)\cdot\exp\left(\lambda(\varphi_Z(0)-1)\right) = \lambda\Delta\mathbb{E}(Z_1) \\
\mathbb{E}\left[(X(t+\Delta) - X(t))^2\right] &= \lambda\Delta\mathbb{E}(Z_1^2) + \lambda^2\Delta^2(E(Z_1))^2
\end{aligned}$$

Simulation of a Compound Poisson process:

For the simulation of a Compound Poisson process we again discretize time. Further we use the fact that for Poisson distributed random variables $N(1) \sim \text{Poisson}(\lambda)$, the time between two consecutive events is exponential distributed with parameter λ, e.g. the time between two events has a cdf of the form $F(x) = 1 - e^{-\lambda x}$. So with the inverse method[6] we can simulate the number of jumps k and the time when the jumps will occur, $0 \leq s_1 < ... < s_k \leq 1$. For a fixed set of equidistant points $0 < t_1 < ... < t_n$, we therefore have just to multiply $s_1, ..., s_k$ with n and round it to the next integer to get the points in the process, where jumps will occur, $t_{\lceil s_i \times n \rceil}$ for $i = 1, ...k$. For each of those points, independent jump-sizes $Z_1, ..., Z_k$ are simulated.

The compound poisson process can be simulated for example in the following way:

$$X(t_{i+1}) = \begin{cases} X(t_i) + p\cdot\text{Gamma}(a_1, b_1) - (1-p)\cdot\text{Gamma}(a_2, b_2) & \text{if a jump occurs at } t_i \\ X(t_i) & \text{else} \end{cases} \tag{2.4.2}$$

Figure 2.4.3 shows 5 paths of a Compound Poisson process with $X(1) - X(0) \sim CP(30, 1, 1, 1, 1, 0.5)$ simulated at a set 100 points with $X(0) = 100$.

2.4.5 Gamma process

Basic distribution:

$$X(t+1) - X(t) \sim \text{Gamma}(a, b)$$

[6] which yields $\frac{-log(U)}{\lambda} = x$

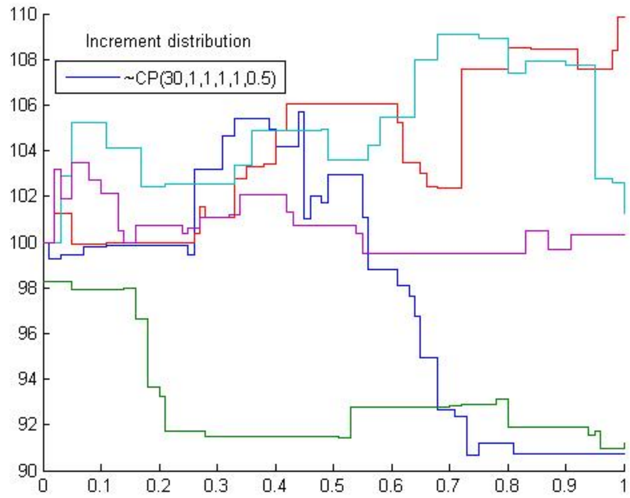


Figure 2.4.3: Compound Poisson process with $\lambda = 30, a_1 = 1, a_2 = 1, b_1 = 1, b_2 = 1$ and $p = 0.5$.

with density

$$f_{a,b}(x) = \frac{1}{\Gamma(a)b^a} x^{a-1} \exp\left(-\frac{x}{b}\right) \qquad x, a, b > 0$$

and characteristic function

$$\varphi(u) = (1 - iub)^{-a}.$$

Distribtution of increment with time interval Δ:

$$X(t+\Delta) - X(t) \sim \text{Gamma}(\Delta a, b)$$

Lévy triplet:

$$\left[a(1 - \exp(-1/b))b, 0, a \cdot \exp(-x/b)x^{-1}\mathbf{1}_{\{x>0)}dx\right]$$

Moments:

$$\begin{aligned}
\mathbb{E}(Y^k) &= \int \frac{1}{\Gamma(a)b^a} x^{a+k-1} \exp\left(-\frac{x}{b}\right) dx \\
&= \frac{\Gamma(a+k)b^k}{\Gamma(a)} \int \frac{1}{\Gamma(a+k)b^{a+k}} x^{a+k-1} \exp\left(-\frac{x}{b}\right) dx \\
&= \frac{\Gamma(a+k)b^k}{\Gamma(a)} = b^k a(a+1)...(a+k-1)
\end{aligned}$$

$$\mathbb{E}Y_t^k = \frac{\Gamma(ta+k)b^k}{\Gamma(ta)}$$

therefore[7] $\mathbb{E}(Y_t) = tab, \mathbb{E}(Y_t^2) = t^2a^2b^2 - tab^2$

Additional facts: In a Gamma process, every path is monotonic.

Simulation of a Gamma process: Euler approximation yields

$$X(t_{i+1}) = X(t_i) + G_{i+1} \text{ for } i = 0, ..., n-1$$

with $G_1, ..., G_n \sim \text{Gamma}(a(t_{i+1} - t_i), b)$ independent resp. $G_1, ..., G_n \sim \text{Gamma}\left(a\frac{1}{n}, b\right)$ if we simulate at equidistant points.

The following figure shows 5 paths of a Gamma process where

$$X(1) - X(0) \sim \text{Gamma}(10, 1)$$

constructed with Euler scheme approximation simulated at a set of 100 points and $X(0) = 100$.

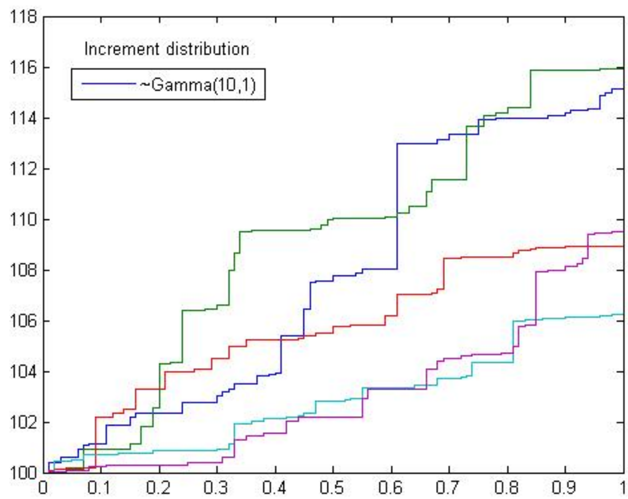


Figure 2.4.4: Gamma process with $a = 10$ and $b = 1$.

2.4.6 Variance Gamma process

If $X_1(\cdot)$ is a Gamma(a, b_1)-process and $X_2(\cdot)$ is an independent Gamma(a, b_2) process, then the difference

$$X(t) = X_1(t) - X_2(t)$$

[7] Notice that $\Gamma(n) = (n-1)!$

is a Variance Gamma process (see [38]) for which we write $\mathrm{VG}(a,b,c)$ with $b = b_1 - b_2$ and $c = b_1 \cdot b_2$.

Increment distribution:

$$X(t+1) - X(t) \sim \mathrm{Gamma}(a, b_1) - \mathrm{Gamma}(a, b_2)$$

The increment distribution of $\mathrm{VG}(a,b,c)$ has characteristic function[8]

$$\varphi(u) = (1 - iub_1)^{-a} \cdot (1 + iub_2)^{-a} = (1 - iub + u^2 c)^{-a}.$$

Distribution of increment with time interval Δ:

$$X(t+\Delta) - X(t) \sim \mathrm{VG}(\Delta a, b, c)$$

with mean[9] $\Delta a(b_1 - b_2) = \Delta ab = \Delta a(G-M)/(MG)$ and variance $\Delta a(b_1^2 + b_2^2) = \Delta a(b^2 + 2c) = \Delta a(G^2 + M^2)(MG)^{-2}$with

$$M = \left(\sqrt{\frac{1}{4}b^2 + c} + \frac{1}{2}b\right)^{-1}$$

$$G = \left(\sqrt{\frac{1}{4}b^2 + c} - \frac{1}{2}b\right)^{-1}$$

Lévy triplet:

$$[\gamma, 0, d\nu(x)]$$

with

$$d\nu(x) = \begin{cases} a \cdot exp(Gx)\frac{1}{|x|}dx & x < 0 \\ a \cdot exp(-Mx)\frac{1}{x}dx & x > 0 \end{cases}$$

and

$$\gamma = \frac{a\left(G(exp(-M) - 1) - M(exp(-G) - 1)\right)}{MG}$$

The VG-process can also be written as a Brownian motion with a drift subjected to a random time, see [38, 10]

$$X(t) = \vartheta Y(t) + \sigma W(Y(t))$$

where $Y(t)$ is a $\mathrm{Gamma}(a,b)$ process with $a = \frac{1}{v} > 0, b = \frac{1}{v} > 0$,$\sigma > 0, v > 0, \vartheta \in \mathbb{R}$.

Simulation of a Variance Gamma process:

$$X(t_{i+1}) = X(t_i) + Z_{i+1} - W_{i+1} \text{ for } i = 0, ..., n-1$$

[8]Notice that $\varphi_{X-Y}(t) = \varphi_X(t) \cdot \varphi_{-Y}(t)$ if X, Y are independent and $\varphi_{-Y}(t) = \varphi_Y(-t)$

[9]Notice that $\mathbb{E}\left[X(t+h) - X(t)\right] = i^{-1}\left[\varphi(u)^h\right]'(0)$

with $Z_1, ..., Z_n \sim \text{Gamma}(a(t_{i+1} - t_i), b_1)$, $W_1, ..., W_n \sim \text{Gamma}(a(t_{i+1} - t_i), b_2)$ independent respectively $Z_1, ..., Z_n \sim \text{Gamma}(a\frac{1}{n}, b_1)$, $W_1, ..., W_n \sim \text{Gamma}(a\frac{1}{n}, b_2)$ if we simulate at equidistant points.

The following figure shows 5 paths of a Variance-Gamma process where $X(1) - X(0) \sim$ Variance-Gamma$(10, 2, 3)$, simulated at a set of 100 points and $X(0) = 100$.

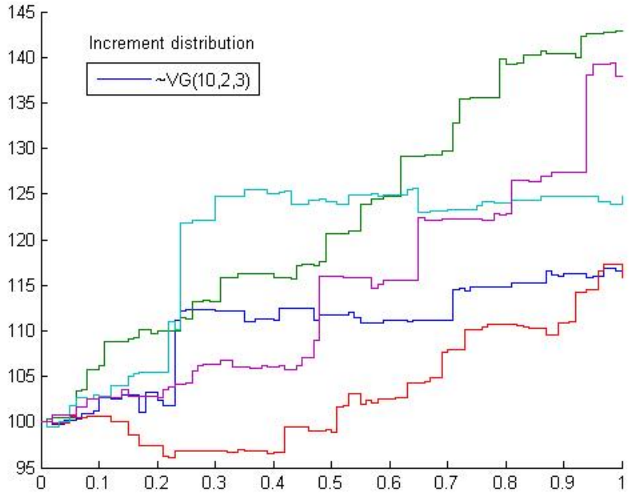


Figure 2.4.5: Variance-Gamma process with $a = 10, b = 2$ and $c = 3$.

2.4.7 Construction of Lévy processes

For details on the construction of Lévy processes see [30, 32]. Considering the Lévy-Chinchin representation, let us look at the characteristic exponent of $X(\cdot)$, $\psi_{X(t)} = \log \varphi_{X(t)}(u)$ with

$$\begin{aligned}
\psi_{X(t)}(u) &= aiut - \frac{1}{2}\sigma^2 u^2 t + t \int\limits_{\mathbb{R}\setminus\{0\}} \left(e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}}\right) d\nu(x) \\
&= \underbrace{aiut - \frac{1}{2}\sigma^2 u^2 t}_{\psi_{X(t)^{(1)}}(u)} + \underbrace{t \int\limits_{|x|\geq 1} \left(e^{iux} - 1\right) d\nu(x)}_{\psi_{X(t)^{(2)}}(u)} + \underbrace{t \int\limits_{|x|<1, x\neq 0} \left(e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}}\right) d\nu(x)}_{\psi_{X(t)^{(3)}}(u)}
\end{aligned}$$

while splitting up the "jumps" into two parts: one part with small jumps $|x| < 1$ and one part with large jumps $|x| \geq 1$.

As we can see, the first part coincides with a Brownian motion with drift and diffusion $X(t)^{(1)} = at + \sigma W_t$. The second part corresponds to a compound Poisson process $X(t)^{(2)} = \sum_{k=1}^{N(t)} Z_k$ with $N(t) \sim \text{Poisson}(\lambda t)$ and $\nu(A) = \lambda P\{Z_1 \in A\}$.
For the third part we use arguments from Martingale theory to show that it converges to a compensated compound Poisson process. $X_t^{(3)} = X_t^{(2)} - \lambda t \int_{|x|<1} x d\mu(x) = X_t^{(2)} - \lambda t \mathbb{E}(Z_1) = X_t^{(2)} - \mathbb{E}(X_t^{(2)})$. Therefore $\mathbb{E}(X_t^{(3)}) = 0$ and so $X_t^{(3)}$ is a martingale.

Remark. Every Lévy process is a semi-martingale [32]

$$X_t = \underbrace{\sigma W_t + X_t^{(3)}}_{martingales} + \underbrace{X_t^{(2)} + at}_{finite\, variation}$$

Remark. The paths of a Lévy process with triplet $[a, \sigma, d\nu(x)]$ have finite variation if $\sigma = 0$ and $\int_{|x|<1} |x| d\nu(x) < \infty$

Remark. The Lévy process with triplet $[a, \sigma, d\nu(x)]$ has finite n-th moments iff

$$\int\limits_{|x|\geq 1} |x|^n d\nu(x) < \infty.$$

3 Lévy processes in finance

In this work, we assume that the price process of the underlying is Markovian. While our methodology applies to any Markov processes, we restrict the presentation and the examples to exponentials of Lévy processes, since these processes are widely used models for stocks prices in Finance.

3.1 Market model

Let us assume that our market consists of two assets. One riskless asset (the bond) with a price $B(t) = B(0)e^{rt}$, and one risky asset (the stock), see [38]. The price process of the underlying risky asset is an *exponential of a Lévy process*, i.e. is modeled as

$$S(t) = S(0) \cdot exp(X(t))$$

with $X = \{X(t), t \geq 0\}$ being a Lévy process with $X(0) = 0$ and $S(0)$ is the today's price. Notice that both processes $X(t)$ and $S(t)$ are Markovian. Further $\log(S(t+\Delta) - \log(S(t))$ follows the distribution of an increment of length Δ of the Lévy process X.
Suppose that a m-dimensional integrable stochastic price vector process

$$S(t) = \left(S(t)^{(1)}, ..., S(t)^{(m)}\right)$$

is given, defined on some probability space $(\Omega, \mathcal{A}, P)$. Each component of $S(t)$ describes the price of one specific asset (stock,bonds etc.). It is assumed that each asset may be traded without transaction costs and with negative holdings (short selling). Typically $S(t) = (S(t), e^{rt})$.

Trading days

The number of trading days in a year is usually assumed to be 252 for stocks. So the life of an option is measured by trading days rather than calendar days and is calculated as T years

$$T = \frac{\text{Number of trading day until option maturity}}{252}$$

see [7, S.271].

3.2 Change of measure

Let $\mathbb{P}$ be the 'real world' measure, which models the price of the underlying $S(t) = S(0)exp(X(t))$. The fundamental theorem of asset pricing states that the absence of arbitrage opportunity is equivalent to the existence of a (not necessarily unique) equivalent measure $\mathbb{Q}$ under which $\tilde{S}(t) = e^{-rt}S(t)$ is a martingale.
Consider any payoff function, f.e. the Vanilla Call Option $g(S(T)) = [S(T) - K]^+$. The fair price $\Pi_0(g(S(T)))$ can be stated in two ways, which are dual to each other:

- $\Pi_0(g(S(T)))$ =minimal initial capital needed to replicate the cashflow $g(S_T)$ by trading $S(t)$ and $B(t)$
- $\Pi_0(g(S(T))) = max_{\mathbb{Q}} \left\{\mathbb{E}_{\mathbb{Q}}\left(e^{-rT}g(S(T))\right)\right\}$

$$s.t. \frac{d\mathbb{Q}}{d\mathbb{P}} > 0,\ e^{-rt}S(t)\ 0 \leq t \leq T \text{ is a martingale under } \mathbb{Q} \tag{3.2.1}$$

The condition that $e^{-rt}S(t)$ is a martingale under $\mathbb{Q}$ implies for all $0 < s < t$ that

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-rt}S(t)|S(0), e^{-rs}S(1), ..., e^{-rs}S(s)\right] = e^{-rs}S(s).$$

In applications, one usually is satisfied to find one martingale measure $\mathbb{Q}$ and set $\Pi_0(g(S(T))) = \mathbb{E}_{\mathbb{Q}}\left(e^{-rT}g(S(T))\right)$. Two important cases have to be distinguished:

- If there exists just one measure $\mathbb{Q}$ which satisifies condition (3.2.1), then the problem reduces to $\Pi_0(g(S(T))) = \mathbb{E}_{\mathbb{Q}}(e^{-rt}g(S(T)))$. This case is called the complete market case and $\mathbb{Q}$ is called the risk-neutral measure.
- If there exist more than one measure $\mathbb{Q}$ which satisifies $\left\{\frac{d\mathbb{Q}}{d\mathbb{P}} > 0, e^{-rt}S(t)\ 0 \leq t \leq T \text{ is a martingale under } \mathbb{Q}\right\}$, there exist infinitely many $\mathbb{Q}$'s because it is a convex set. This case is referred to incomplete market.

With this risk-neutral probabilities, every asset can be priced by simply taking its expected payoff (i.e. assets are priced as if investors were risk neutral). If we used the physical probabilities, every asset would require a different adjustment, because they differ in riskiness. Under the assumption of *No Arbitrage* one can first adjust the probabilities such that they incorporate the effects of risk, and then take the expectation under this new probability distribution.
So our model assumptions are:

- No Arbitrage: There is no opportunity to get a risk-free profit. The no-arbitrage condition is equal to requiring the existence of the risk neutral measure $\mathbb{Q}$ under which
$$\Pi_0(g(S(T))) = \mathbb{E}_{\mathbb{Q}}(e^{-rt}g(S(T)))$$
- Every possible derivative can be hedged. We can construct a portfolio consisting of a risk free cash bond and underlying stocks which has exactly the value of the derivative at maturity.

In applications, one usually is satisfied to find just one martingale measure $\mathbb{Q}$. For exponentials of Lévy processes an equivalent martingale measure can be obtained for example by the Esscher transform or mean-correction, which just changes the parameter values of the original Lévy process, see [38].

The Esscher Transform Let $\mathbb{P}$ be the 'real world' measure, which models the price of the underlying $S(t) = S(0)exp(X(t))$. Assume $S(0) = 1$. Our aim is to find a $\mathbb{Q}$ under which $\tilde{S}(t) = e^{-rt}S(t)$ is a martingale. Via an exponential transform, we can define a new density

$$\frac{d\mathbb{Q}^\vartheta}{d\mathbb{P}} = \frac{exp(\vartheta X(T))}{\mathbb{E}_\mathbb{P}(exp(\vartheta X(T)))}$$

Now we try to find a ϑ^*such that $e^{-rt}S(t)$ is a martingale under $\mathbb{Q}$. From

$$\mathbb{E}^{\vartheta^*}\left(e^{-r(t+h)}S(t+h)|S(t)\right) = \mathbb{E}^{\vartheta^*}\left(e^{-rt}e^{-rh}S(t)exp(\underbrace{X(t+h)-X(t)}_{S(t+h)/S(t)})|S(t)\right) = e^{-rt}S(t)$$

we see that a necessary condition is

$$1 = e^{-rh}\mathbb{E}^{\vartheta^*}(exp(X(h)))\ \forall h.$$

To derive ϑ^*out of that, we do some calculations.

$$\begin{aligned} 1 &= exp(-rt)\mathbb{E}_\mathbb{P}\frac{exp(X(t))\cdot exp(\vartheta^* X(t))}{\mathbb{E}(exp(\vartheta^* X(t)))} \\ 1 &= exp(-rt)\frac{\mathbb{E}_\mathbb{P}(exp(\vartheta^*+1)X(t)))}{\mathbb{E}_\mathbb{P}(exp(\vartheta^* X(t)))} \end{aligned}$$

Let $u = -i\vartheta^*$ and therefore $\varphi(-i\vartheta^*) = \mathbb{E}_\mathbb{P}(exp(\vartheta^* X(1)))$.
Consider also that $[\varphi(u)]^t = \mathbb{E}_\mathbb{P}(exp(iuX(t)))$. So we get

$$1 = exp(-rt)\frac{\varphi^t(-i(\vartheta^*+1))}{\varphi^t(-i\vartheta^*)}.$$

Set $t = 1$

$$\begin{aligned} 1 &= exp(-r)\frac{\varphi(-i(\vartheta^*+1))}{\varphi(-i\vartheta^*)} \\ exp(r) &= \frac{\varphi(-i(\vartheta^*+1))}{\varphi(-i\vartheta^*)} \end{aligned} \tag{3.2.2}$$

Via equation 3.2.2 we calculate ϑ^*. Then $e^{-rt}S(t)$ is a martingale under $\mathbb{Q}^{\vartheta^*}$ with $\frac{d\mathbb{Q}^{\vartheta^*}}{d\mathbb{P}} = \frac{exp(\vartheta^* X(T))}{\mathbb{E}_\mathbb{P}(exp(\vartheta^* X(T)))}$.

If we have found ϑ^*, another question arises: How does the Lévy triplet change if we use

the "equivalent measure"?
The answer is you have to correct the triplet using following transformation:

$$\begin{aligned}
\varphi^{(\vartheta^*)}(u) &= \mathbb{E}_{\mathbb{Q}^{\vartheta^*}}\left(exp(iuX(1))\right) \\
&= \mathbb{E}_{\mathbb{P}}\left(exp(iuX(1))\frac{exp(\vartheta^* X(1))}{\mathbb{E}(exp(\vartheta^* X(1)))}\right) \\
&= \frac{\varphi(u-i\vartheta^*)}{\varphi(-i\vartheta^*)}
\end{aligned}$$

So

$$\begin{aligned}
\psi^{(\vartheta^*)}(u) &= log\varphi^{(\vartheta^*)} = log\varphi(u-i\vartheta^*) - log\varphi(-i\vartheta^*) \\
&= i\gamma(u-i\vartheta^*) - \frac{1}{2}\sigma^2(u^2 - 2iu\vartheta^* - {\vartheta^*}^2) \\
&+ \int_{-\infty}^{\infty}\left(e^{(iu+\vartheta^*)x} - 1 - (iu+\vartheta^*)x\mathbf{1}_{\{|x|<1\}}\right)d\nu(x) \\
&- \left(\gamma\vartheta^* + \frac{1}{2}\sigma^2{\vartheta^*}^2 + \int_{-\infty}^{\infty}\left(e^{\vartheta^* x} - 1 - \vartheta^* x\mathbf{1}_{\{|x|<1\}}\right)d\nu(x)\right) \\
&= iu(\gamma+\vartheta^*\sigma^2) - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{\infty}\left(e^{\vartheta^* x}(e^{iux}-1) - iux\mathbf{1}_{\{|x|<1\}}\right)d\nu(x)
\end{aligned}$$

The Lévy triplet under the Esscher transform is

$$\left[\gamma^{(\vartheta^*)}, \sigma^{(\vartheta^*)}, d\nu^{(\vartheta^*)}(x)\right]$$

with

$$\begin{aligned}
\gamma^{(\vartheta^*)} &= \gamma + \sigma^2\vartheta^* + \int_{-1}^{1}\left(exp(\vartheta^* x) - 1\right)d\nu(x) \\
{\sigma^{(\vartheta^*)}}^2 &= \sigma^2 \\
d\nu^{(\vartheta^*)}(x) &= exp(\vartheta^* x)d\nu(x)
\end{aligned}$$

Example ($X_1 \sim$ Poisson(λ))**.** :

The characteristic function $\varphi(u) = \mathbb{E}[exp(uiX_1]$ is given by

$$\varphi(u) = exp(\lambda(e^{iu}-1)).$$

So for ϑ^*do

$$\begin{aligned}
exp(r) &= \frac{\varphi(-i(\vartheta^*+1))}{\varphi(-i\vartheta^*)} \\
&= \frac{exp(\lambda(e^{\vartheta^*+1}-1))}{exp(\lambda(e^{\vartheta^*}-1))} \\
&= exp\left(\lambda(e^{\vartheta^*+1}-e^{\vartheta^*})\right) \\
r &= \lambda(e^{\vartheta^*+1}-e^{\vartheta^*}) \\
log(r/\lambda) &= log(e^{\vartheta^*}(e^1-1)) \\
log(r/\lambda) &= loge^{\vartheta^*}+log(e^1-1) \\
log(r/\lambda)-log(e-1) &= \vartheta^*
\end{aligned}$$

So under $\mathbb{Q}^{\vartheta^*}$, X_1 has the Lévy triplet:

$$\begin{aligned}
\gamma^{\vartheta^*} &= \underbrace{\sigma^2\left(log(r/\lambda)-log(e-1)\right)}_{=0}+\underbrace{\int_{-1}^{1}\left(e^{(log(r/\lambda)-log(e-1))x}-1\right)\lambda\delta_1}_{=0}=0 \\
\sigma^{\vartheta^*2} &= \sigma^2=0 \\
d\nu^{\vartheta^*}(x) &= e^{(log(r/\lambda)-log(e-1))x}\lambda\delta_1
\end{aligned}$$

This leads to

$$\begin{aligned}
\psi(u) &= \int_{-\infty}^{\infty}(e^{iux}-1-iux\mathbf{1}_{\{|x|<1\}})e^{(log(r/\lambda)-log(e-1))x}\lambda\delta_1 \\
&= \lambda(e^{iu}-1)(r/\lambda)\frac{1}{e-1} \\
&= \frac{r}{e-1}(e^{iu}-1)
\end{aligned}$$

The Poisson process under equivalent martingale measure is given by:

$$S(t)=S(0)e^{P\left(\frac{r}{e-1}t\right)}$$

The Mean-Correcting Martingale measure Usually the Esscher Transform is very easy to apply. But for some cases (if the market is incomplete), the equivalent martingale measure has to be obtained in a different way, for example by mean correcting the exponential of a Lévy process, see [38]. For this, an additional drift parameter m will be introduced, which play a special role in risk neutral modelling. The stock price is then described by

$$S(t)=S(0)\exp(mt+X(t)).$$

For the Mean-Correcting Martingale measure we proceed as follows. After estimating all the parameters that are involved in the process, we change the paramter m such that the discounted stock price becomes a martingale. The change of paramter m is done in the following way

$$m_{new} = m_{old} + r - \log \varphi(-i)$$

where $\varphi(x)$ is the characteristic function of $X(t)$ including the parameter m_{old}. The discounted stock price $\tilde{S}(t) = e^{-rt} S(t)^{new}$ with $S(t)^{new} = S(0)\exp(m_{new}t + X(t))$ is then a martingale.

Example. Black-Scholes model

$$S(t) = S(0)\exp(mt + X(t))$$

with $X(t) \sim N\left(0, \sigma^2 t\right)$ and $m = \mu - \frac{1}{2}\sigma^2$. Therfore

$$m_{new} = \mu - \frac{1}{2}\sigma^2 + r - \mu = r - \frac{1}{2}\sigma^2$$

and $\tilde{S}(t) = e^{-rt} S(0) e^{(r - \frac{1}{2}\sigma^2) + X(t)}$ is a martingale.

3.3 Exponential Lévy processes

We now turn to some relevant examples of Lévy-driven stock models which are typically used in finance.

3.3.1 Geometric Brownian Motion model

The Geometric Brownian motion model (GBM model) -or Black-Scholes model- is the most commonly used model and can be found in detail in [12]. A stochastic process $S(t)$ is a *Geometric Brownian motion* if log$S(t)$ follows a Brownian motion with inital value log$S(0)$. It was basically the work of Paul Samuelson in the 1960s to use GBM as a model in finance.
The driving process is a Brownian motion with constant drift μ and constant diffusion σ^2

$$X(t) = \mu t + \sigma W(t)$$

where $W(t)$ is a standard Brownian motion. According to Itô's formula, the stock prices follows the process

$$S(t) = S(0)\exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma dW(t)\right]$$

and more general for $u < t$:

$$S(t) = S(u)exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t-u) + \sigma(W(t) - W(u))\right)$$

Simulation of Geometric Brownian motion For a fixed set of equidistant points $0 < t_1 < ... < t_n$[1]:

$$\begin{aligned} S(t_{i+1}) &= S(t_i)exp\left([\mu - \frac{1}{2}\sigma^2](t_{i+1}-t_i) + \sigma\sqrt{t_{i+1}-t_i}Z_{i+1}\right) \\ &= S(t_i)exp\left([\mu - \frac{1}{2}\sigma^2]\left(\frac{1}{n}\right) + \sigma\sqrt{\frac{1}{n}}Z_{i+1}\right) \qquad i = 0, 1, ..., n-1 \end{aligned}$$

with $Z_1, Z_2, ..., Z_n$ are indenpendent standard normals [12].

Risk neutral measure:

With the help of the Esscher transform (see paragraph 3.2) we want to find $\mathbb{Q}^{\vartheta^*}$ under which $e^{-rt}S(t)$ is a martingale. The characteristic function $\varphi(u) = \mathbb{E}[\exp(uiX(1)]$ is given by $\varphi(u) = \exp\left(iu(\mu - \frac{1}{2}\sigma^2) - \frac{1}{2}\sigma^2u^2\right)$. So for ϑ^* we have to calculate

$$\begin{aligned} \exp(r) &= \frac{\varphi(-i(\vartheta^*+1))}{\varphi(-i\vartheta^*)} \\ \exp(r) &= \exp\left((\vartheta^*+1)(\mu - \frac{\sigma^2}{2}) + \frac{\sigma^2(\vartheta^*+1)^2}{2} - \vartheta^*(\mu - \frac{\sigma^2}{2})\frac{\sigma^2{\vartheta^*}^2}{2}\right) \\ &= \exp\left(\mu - \frac{\sigma^2}{2} + \sigma^2\vartheta^* + \frac{\sigma^2}{2}\right) \\ r &= \mu + \sigma^2\vartheta^* \end{aligned}$$

and therefore

$$\vartheta^* = \frac{r-\mu}{\sigma^2}.$$

So under $\mathbb{Q}^{\vartheta^*}$, $X(1)$ has the Lévy triplet

$$\left[r - \frac{\sigma^2}{2}, \sigma^2, 0\right].$$

Therefore

$$\psi(u) = iu(r - \frac{1}{2}\sigma^2) - \frac{1}{2}\sigma^2u^2$$

and $\tilde{S}(t) = e^{-rt}S(0)e^{(r-\frac{1}{2}\sigma^2)+\sigma W(t)}$ is a martingale.

[1] we set $t_{i+1} - t_i = \frac{1}{n}$

3.3.2 Variance Gamma model

The driving process is a Variance Gamma process with increment distribution

$$X(t+1) - X(t) \sim \mathrm{VG}(a, b, c).$$

The stock prices follow the process

$$S(t) = S(0)\exp(\mathrm{VG}(at, b, c))$$

Simulaton of exponential Variance Gamma process

$$\begin{aligned} S(t_{i+1}) &= S(t_i)e^{(\mathrm{Gamma}((t_{i+1}-t_i)a,b_1)-\mathrm{Gamma}((t_{i+1}-t_i)a,b_2))} \\ &= S(t_i)e^{\left(\mathrm{Gamma}\left(\frac{1}{n}a,b_1\right)-\mathrm{Gamma}\left(\frac{1}{n}a,b_2\right)\right)} \quad \forall i = 0, ..., n-1 \end{aligned}$$

Risk neutral measure Under all exponential Lévy processes, only those who are multiples of a Brownian motion with drift or a multiple of the Poisson process with drift result in a complete Lévy market, see [43]. The market modeled on the exponential VG process is an incomplete market. Therefore the risk-neutral measure is not unique and a choice of measure has to be done. Recall that the characteristic function of a VG-process is of the form $\varphi(u) = \left(1 - iub + u^2c\right)^{-a}$. According to the mean-correcting martingale measure, the drift term m such that $S(t) = S(0)\exp(m + X(t))$ is a martingale has to be[2]

$$\begin{aligned} m &= \underbrace{m_{old}}_{0} + r - \log\varphi(-i) \\ &= r + a\log(1 - b - c) \end{aligned}$$

or

$$m = r + C\log((M-1)(G+1)/(MG))$$

see [38], so $e^{-rt}S(t) = e^{-rt}S(0)\exp(m + X(t))$ is a martingale.
The parameters of the model can be estimated through market data. In [38] this happens by minimizing the root-mean-square error between the markets and model's price.

3.3.3 Compound Poisson model

The model for the risky asset $S(t) = S(0)exp(X(t))$ is said to be a Compound Poisson model, if $X(t)$ follows a compound Poisson process

$$X(t) = \sum_{k}^{N(t)} Z_k \ \ t \geq 0.$$

[2]Notice: First of all estimate the parameters of the process under $m = 0$. Then introduce m as a function of the estimated parameters.

In that case we write

$$X(\cdot) \quad \sim \quad \mathrm{CP}(\lambda, a1, b1, a2, b2, p)$$

Simulation of the exponential Compound Poisson process For simulating the Compound Poisson process, we derive the points where jump occurs as described in subsection 2.4.4 on page 18. In a simulation we have

$$S(t_{i+1}) = \begin{cases} S(t_i)e^{p\cdot \mathrm{Gamma}(a_1,b_1)-(1-p)\cdot \mathrm{Gamma}(a_2,b_2)} & \text{if a jump occurs at } t_i \\ S(t_i) & \text{else} \end{cases}$$

Risk neutral measure: For the risk neutral measure of a Compound Poisson process model let us first consider the case with $p = 1$. Then the characteristic function of the Compound Poisson process is

$$\varphi(u) = \exp\left(\lambda\left(\varphi_z(u) - 1\right)\right)$$

with $\varphi_Z(u) = (1 - iub)^{-a}$. The discounted stock $e^{-rt}S(t) = e^{-rt}S(0)\exp(m + X(t))$ is a martingale, if the drift term m is of the form

$$\begin{aligned} m &= \underbrace{m_{old}}_{0} + r - \log\varphi(-i) \\ &= r - \lambda\left((1-b)^{-a} - 1\right). \end{aligned}$$

For the case $0 < p < 1$ consider the characteristic function

$$\varphi_Z(u) = p(1 - iub_1)^{-a_1} + (1-p)(1 + iub_2)^{-a_2}.$$

So for $e^{-rt}S(t) = e^{-rt}S(0)\exp(m + X(t))$ to be a martingale, m has to be

$$\begin{aligned} m &= \underbrace{m_{old}}_{0} + r - \log\varphi(-i) \\ &= r - \lambda\left(p(1-b_1)^{-a_1} + (1-p)(1+b_2)^{-a_2} - 1\right). \end{aligned}$$

4 Sensitivity (Gradient) estimation methods

In this section, the basic ideas of gradient estimation techniques are introduced. The focus is on Measure valued differentiation.

4.1 Basic Problem

Let $S_\theta(t)$ be a stochastic process on $(\Omega, \mathcal{F}, \nu_\theta)$, where θ is a real parameter. Furthermore let $S_\theta(t)$ be a Markovian process, with a transition law dependent on the parameter θ. For any functional $\mathcal{A}[S_\theta(.)]$ we are interested in the sensitivity w.r.t. the parameter θ, i.e. the gradient

$$\frac{\partial}{\partial \theta}\mathbb{E}\left[\mathcal{A}\left(S_\theta(.)\right)\right]$$

The functional $\mathcal{A}\left(S_\theta(.)\right)$ may be

- $\mathcal{A}\left(S_\theta(.)\right) = H(S_\theta(T))$
- $\mathcal{A}\left(S_\theta(.)\right) = H(S_\theta(\infty))$
- $\mathcal{A}\left(S_\theta(.)\right) = \int_0^T H(S_\theta(t))dt$
- $\mathcal{A}\left(S_\theta(.)\right) = \int_0^\tau H(S_\theta(t))dt$, where τ is some stopping time.

EXAMPLES:

- Markov Systems (queueing, service, manufacturing)
  Suppose that θ denotes the parameters of a Markov System (queueing, inventory, renewal). Let $S_\theta(t)$ be the state of the system at time t and $S_\theta(\infty)$ the steady state (if exists). Then
  - $\mathbb{E}[H(S_\theta(T))]$is the performance of the system at time T
  - $\mathbb{E}[\int H(S_\theta(t))]dt$ is the expected integrated transient behavior
  - $\mathbb{E}[H(S_\theta(\infty))]$ is the expected stationary behavior.
- Finance
  Let $S_\theta(t)$ describe the evolution of an underlying asset, for a parameter vector θ. Let H be the payoff function of a contingent contract. Then $\mathbb{E}[H(S_\theta(T))]$ is the value of the contingent contract (European type). If τ is a stopping time, then $\mathbb{E}[H(S_\theta(\tau))]$ is the value of the American type contingent contract. One wants to estimate the sensitivity of the price w.r.t. the parameter θ.

We are interested in finding the value of

$$\frac{\partial}{\partial\theta}\mathbb{E}\left[\mathcal{A}\left(S_\theta(.)\right)\right]$$

and -more generally-

$$\frac{\partial^k}{\partial\theta^k}\mathbb{E}\left[\mathcal{A}\left(S_\theta(.)\right)\right]$$

in a way to estimate it based on sampling. Our attention further will be on Finance applications. So our aim is to calculate

$$\frac{\partial}{\partial\theta}\mathbb{E}_Q[e^{-rT}H(S_\theta(T))]. \tag{4.1.1}$$

Since closed form expressions do not exist (except for very special cases), estimations of (4.1.1) with high accuracy are needed. There exist several Monte Carlo techniques for estimating the Greeks.
The simpliest form is the *finite difference (FD) approximation,* which is in general biased and its use requires balancing bias against variance. More sophisticated unbiased methods are subdvided in methods which assume parametrized integrands

$$\theta \mapsto \int H(\theta,\omega)d\nu(\omega)$$

and methods with parametrized integrators

$$\theta \mapsto \int H(\omega)d\nu_\theta(\omega)$$

see [33]. The *pathwise method (Infinitesimal Pertubation Analysis IPA)* is a method of the first sort assuming functional dependence of S on θ. To ensure unbiasedness, the IPA method requires the payoff function to be Lipschitz continuous among several other conditions which are required, see [12]. Other direct methods belong to the second group as the *Likelihood ratio method* or *Score Function method*, [12], the *Malliavin calculus*, [37, 26, 39], and the *Measure valued differentiation*, [3, 4, 5, 18, 41].

4.2 Finite difference (FD) method

An obvious (and easy to implement) approach to estimate derivatives, is the *finite difference (FD) approximation*, see [12].

Let us use F to denote the discounted payoff of an option depending on θ, the parameter of interest. In the case of a Call Option we have $F(\theta) = e^{-rT}[S_\theta(T) - K]^+$. Notice that in applications, $F(\theta)$ is the discounted payoff of an option at time T, and $\mathbb{E}[F(\theta)]$ is its price.θ can be for example the initial price, the volatility, the maturity time, the risk-free interest rate... . Further let us assume that we can simulate $F(\theta)$ for every value of θ. The question is how to estimate

$$\frac{d}{d\theta}\mathbb{E}[F(\theta)].$$

Let $\beta(\theta) = \mathbb{E}[F(\theta)]$, then the problem is to estimate $\beta'(\theta)$.
According to the *forward FD method,* see[12, 20], the first order Greek is approximated by

$$\hat{\Delta}_{FD} = \frac{\bar{F}_n(\theta + h) - \bar{F}_n(\theta)}{h}$$

for some $h > 0$ where $\bar{F}_n(\theta + h)$ and $\bar{F}_n(\theta)$ denote the average of n replications of the model at parameter $(\theta + h)$ respectively θ. A more accurate but more costly method would be the *centred FD method*, see [12].
If β is twice differentiable at θ one can see that the bias is

$$Bias\hat{\Delta}_{FD} = E[\hat{\Delta}_{FD} - \beta'(\theta)] = \frac{1}{2}\beta''(\theta)h + o(h^2)$$

If $F(\theta)$ is differentiable, the FD approximation converge to the true value if h is small enough but numerically this is not the case. Because the variance of the FD estimator is

$$Var[\hat{\Delta}_{FD}] = h^{-2}Var[\bar{F}_n(\theta + h) - \bar{F}_n(\theta)]$$

one can see that if h is very small, the esitmate will get unstable. Even with variance reduction techniques the problem still remains and especially higher-order Greek estimates are in most cases numerically unstable. Although the FD method is easy to understand and implement, the weakness is to decide on the right h. An additional problem lies in the compuational effort, because of the extra simulation of n additional replicants, [22].

4.3 Pathwise method

The *pathwise method*, or *infinitesimal perturbation analysis (IPA)* is a direct method, and therefore unbiased. It extracts, if possible, the parameter of interest out of the distribution and push it into the functional H. Let H be a payoff function. The pathwise method assumes a functional dependence of H on θ.
If there exists a process representation (defined later) of ν_θ, the pathwise method transforms $\theta \mapsto \int H(\omega)d\nu_\theta(\omega)$ to the form $\theta \mapsto \int H(\theta,\omega)d\nu(\omega)$.

Example. $\nu_\theta : \theta \mapsto N(\theta, 1)$ familiy of distributions

$$\mathbb{E}_\theta[H(S)] = \int H(\omega)d\nu_\theta(\omega)$$

Under Transformation, ν_θ can be represented as: $S_\theta = \theta + Z$ with $Z \sim N(0,1) = \nu$. The distribution we have to sample from, is no longer dependent on the parameter θ and we have

$$\mathbb{E}[H(S_\theta)] = \int H(\theta,\omega)d\nu(\omega).$$

To get the derivative $\frac{\partial \mathbb{E}[H(S_\theta)]}{\partial \theta}$, the IPA differentiates the payoff function w.r.t. S_θ, and multiplies this with the inner derivative $\frac{\partial S_\theta}{\partial \theta}$.

The following definitions can be found in [33].

Definition. Let (ν_θ) be some family of probability measures. A collection of random variables S_θ is called a process representation for (ν_θ), if S_θ has distribution ν_θ.

The *process representation* S_θ can be computed for example with the inverse transformation method.

Definition. The family (ν_θ) of probability measures is called process p-differentiable at point θ_0, if there it has (in a neighbourhood of θ_0) a process representation (S_θ) for (ν_θ), such that $\theta \mapsto S_\theta$ is L^p-differentiable at θ_0.

That means that if ν_θ is L^p-differentiable at point θ, there is a random variable Z_θ such that

$$||s||^{-1}||S_{\theta+s} - S_\theta - sZ_\theta||_p \to 0 \text{ as } s \to 0.$$

The random variable Z_θ is the *process derivative* of (ν_θ) at point θ and (S_θ, Z_θ) is called the *process derivative pair* .

Lemma (Fundamental property of process derivatives)**.** *Suppose that $H(s)$ is Lipschitz continuous, and its derivative exists with probability* 1*. Suppose further that $\theta \mapsto \nu_\theta$ is process* 2*-differentiable with derivative pair(S_θ, Z_θ). Then*

$$\nabla_\theta \int H(\omega) d\nu_\theta(\omega) = \mathbb{E}(Z_\theta H'(S_\theta))$$

see [33].

In the following we restrict our attention to payoff functions of the form

$$F(\theta) = H(S_1(\theta), ..., S_m(\theta)),$$

where F depends on the underlying assets at a finite number of fixed dates. We assume functional dependence of S (and therefore H) on a parameter θ. In the following we consider the one-dimensional case, where H is a real valued mapping such that $\mathbb{E}[H(S_\theta)]$ is defined for any $\theta \in (0, \infty)$. The goal is to calculate

$$\frac{d}{d\theta}\mathbb{E}[H(S_\theta)].$$

If $H(S_\theta)$ is differentiable and the interchange of expectation and differentiation is justified, the derivative can be estimated in a straightforward way by

$$\frac{d}{d\theta}\mathbb{E}[H(S_\theta)] = \mathbb{E}\left[\frac{d}{d\theta}H(S_\theta)\right] = \mathbb{E}\left[\left(\frac{d}{d\theta}S_\theta\right) H'(S_\theta)\right]$$

where $H'(S_\theta)$ denotes the derivative of $H(S_\theta)$ with respect to S_θ, see [24, 40, 21, 28]. The derivative $\frac{d}{d\theta}S_\theta$ is called the process derivative of S_θ.

Example (Black-Scholes model)**.** Our aim is to compute the Δ of a Vanilla Call Option[1]. The process representation of the stock price in the Black-Scholes case is

$$S(T) = S(0)e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z}, \qquad Z \sim N(0,1).$$

Let F describe the discounted payoff of a Vanilla Call Option at time T, so

$$F(\theta) = H(S_\theta(T)) = e^{-rT}[S_\theta(T) - K]^+.$$

To calculate the options delta, we have to differentiate $\mathbb{E}[F(\theta)]$ with respect to $S(0)$. Recall

$$\frac{d}{d\theta}\mathbb{E}[F(\theta)] = \frac{d}{d\theta}\mathbb{E}[H(S_\theta(T))] = \mathbb{E}\left[\frac{d}{d\theta}H(S_\theta(T))\right] = \mathbb{E}\left[\left(\frac{d}{d\theta}S_\theta(T)\right)H'(S_\theta(T))\right].$$

So for the Δ of a Vanilla Call Option we have to calculate

$$\frac{dF}{dS(0)} = \underbrace{\frac{dH(S_{S(0)}(T))}{dS_{S(0)}(T)}}_{H'(S_{S(0)}(T))} \underbrace{\frac{dS_{S(0)}(T)}{dS(0)}}_{\text{process derivative}}$$

from which we get

$$\frac{dF}{dS(0)} = e^{-rT}\frac{S(T)}{S(0)}\mathbf{1}\{S(T) > K\},$$

which can be easily computed in simulation and the expected value of this estimator is the Black Scholes Delta.

At this point we have to mention the weaknesses of the pathwise method. It is not always possible to differentiate sample paths or to interchange expectation and differentiation without of loss of unbiasedness of the estimator. Consider for example the options gamma of the previous example or a Digital Option with $F = e^{-rt}1\{S(T) > K\}$. It follows that

$$\mathbb{E}\left[\frac{dF}{dS(0)}\right] = 0 \neq \frac{d}{dS(0)}\mathbb{E}[F].$$

The payoff function here is not continuous as $S(T)$ crosses the strike. To guarantee the unbiasedness of the IPA estimator a necessary and sufficient condition is uniform integrability of the difference quotients $h^{-1}[F(\theta + h) - F(\theta)]$, see [12], but in praxis this condition is very bulky.
More practicable conditions are that the process representation has to exist, the process derivative S'_θ has to exist with probability one at each θ. Further

$$F'(\theta) = \frac{\partial H}{\partial S_\theta}(S_\theta)S'_i(\theta)$$

[1]The delta of an option is its derivative with respect to the initial value $S(0)$

has to exist with probability one. The payoff function $H(S_\theta)$ has to be Lipschitz continuous with probability one , and its Lipschitz constant should have finite first moments. For further details see [12, 40].
We summarize: if a process representation exist, and if its derivative, the process derivative exist with probability 1 and if the payoff function is lipschitz continuous, than the interchange of expectation and differentitation is justified and we may compute the derivative in a straightforward way via the IPA method. We see there are many requirements for the applicability of the pathwise method. Anyhow this method produces an unbiased, low variance estimator. The requirement of Lipschitz continuity is a very strong restriction and typically fails to hold especially for Rainbow Options, Digital Options or when calculating the gamma of ordinary Call Options.

The following example shows how IPA applies to stochastic differential equation models.

Example (SDE: BS-model)**.** Let us consider a SDE of the form

$$dS_r(t) \quad = \quad f_r(S_r(t))dt + \sigma_r(S_r(t))dW(t)$$

with $0 < t_1 < ... < t_n$. In the Black Scholes framework we have

$$dS_r(t) = \left(r - \frac{1}{2}\sigma^2\right) S_r(t)dt + \sigma S_r(t)dW(t)$$

For estimating[2] the ρ, we can proceed as described in 2.3. Let $\frac{dS_r(t)}{dr} = D_r(t)$, then we have

$$D_r(t+h) \approx D_r(t) + \left[\dot{f}_r(S_r(t)) + f_r'(S_r(t))D_r(t)\right] h + \left[\dot{\sigma}_r(S_r(t)) + \sigma_r'(S_r(t))D_r(t)\right] \sqrt{h}Z$$

and therefore

$$D_r(t+h) \approx D_r(t) + \left[S_r(t) + (r - \frac{1}{2}\sigma^2)D_r(t)\right] h + [\sigma D_r(t)] \sqrt{h}Z.$$

Considering the ρ for a Vanilla Call Option, recall the payoff function $H(S(T)) = e^{-rT}[S(T) - K]^+$:

$$\begin{aligned} \frac{\partial \mathbb{E}\left(H(S(T))\right)}{\partial r} &= \mathbb{E}\left(-Te^{-rT}[S(T) - K]^+ + H'(S(T))D_r(T)\right) \\ &= \mathbb{E}\left(-Te^{-rT}[S(T) - K]^+ + e^{-rt}\mathbf{1}_{\{S_T > K\}}D_r(T)\right). \end{aligned}$$

In the classical BS-framework there exists a exact solution, $D_r(T) = t \cdot S(T)$. So we do not have to simulate the process $D_r(t)$ which takes additional time. But in the case of more complicated functions $f(.)$ and $\sigma(.)$, this is the way how to proceed.

[2] ρ is the sensitivity of the payoff function with respect to the interest rate r

4.4 Score-Function method

The *Score Function method* or *Likelihood ratio method (LRM)* , see [12, 24, 28], is another direct method. Different to the pathwise method, the score function method regards the parameter θ as a distributional parameter. The interchange of expectation and differentiation is here not such a big deal because probabilities are typically smooth functions of their parameters whereas option payoffs are not.

Suppose S has a probability density f with distribution parameter θ. For the expectation we therefore write

$$\mathbb{E}_\theta[H(S)] = \int H(s) f_\theta(s) ds.$$

Assuming that the interchange of expecation and differentiation is justified and that f_θ is differentiable. Therefore

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(S)] = \int H(s) \frac{d}{d\theta} f_\theta(s) ds.$$

The problem is that $\frac{d}{d\theta} f_\theta(s)$ is not a density and therefore we cannot sample from it. To come up with this problem we multiply and divide the integrand by f_θ. This yields

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(S)] = \int H(s) \frac{\frac{d}{d\theta} f_\theta(s)}{f_\theta(s)} f_\theta(s) ds = \mathbb{E}_\theta\left[H(S)\frac{\dot{f}_\theta(S)}{f_\theta(S)}\right].$$

Note that

$$\frac{\dot{f}_\theta(S)}{f_\theta(S)} = \frac{d}{d\theta}\log(f_\theta(s)).$$

Let

$$SF(\theta, s) = \frac{d}{d\theta}\log(f_\theta(s))$$

then $SF(\theta, x)$ is the socalled *Score Function* and the derivative can be written as

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(S)] = \mathbb{E}_\theta[H(S) SF(\theta, S)].$$

The Likelihood ratio method leads to an unbiased estimator. Different to the IPA method, the payoff function here is not required to be Lipschitz continuous. One problem of the Score function method is the need of knowledge of the explicit density and its differentiability. Another disadvantage is that the LRM often produces high variance, especially if we want to estimate sensitivites of a path dependent payoff function w.r.t. a parameter that occurs at every step of the path. Summing a large number of terms in the score produces high variance, for further details see [12].

4.5 Malliavin Calculus

This section gives a brief informal introduction to Malliavin calculus and its appliance to the estimation of the Greeks. We therefore introduce two basic operators on which Malliavin calculus relies, the *Malliavin Derivative* and its adjoint operator (*Skorohod integral*), see [37, 9, 11]. Further we will discuss how to derive Greeks with the help of Malliavin calculus via an integration by parts procedure, wich can be found in detail in [26].

4.5.1 Introduction

Malliavin calculus is a stochastic calculus of variations of stochastic processes and can be applied for example to random variables which depend on trajectories of Brownian motion. Such random variables we call from now on *Brownian Functionals.*
Let $(t_1, ..., t_n)$ be a partition of $[0, T]$, $0 \leq t_1 < ... < t_n \leq T$ and let $W = (W_{t_1}, ..., W_{t_n})$ be a sample of a Brownian motion at those points. We than can construct a random variable

$$F(W) \equiv f(W(t_1), ..., W(t_n))$$

whereas f is a function which is infintely differentiable with derivatives of polynomial growth, so $f \in C_p^\infty(\mathbb{R}^n)$. $F(W)$ is called a *Smooth Brownian Functional,* i.e. $F \in S_p^\infty(\mathbb{R}^n)$.

Example (Brownian Functionals)**.**

- Quadratic function: $F(W) = \sum_{j=1}^n W(t_j)^2$
- Polynomials: $F(W) = \sum_{k=1}^K a_k W(T)^k$
- Stock price in Black Scholes model: $S(T) = f(W(T)) = S(0)e^{(r-\frac{1}{2}\sigma^2)T+\sigma W(T)}$

4.5.2 Malliavin Derivative

The Malliavin Derivative measures the effect of a small change in the trajectory of the Brownian motion on the value of $F(W)$.

Definition (Malliavin Derivative)**.** The Malliavin Derivative of F at t is defined by

$$D_t F(W) = \sum_{i=1}^n \frac{\partial F}{\partial x_i}(W(t_1), ..., W(t_k), ..., W(t_n)) \mathbf{1}_{[t,\infty[}(t_i)$$

where $\frac{\partial F}{\partial x_i}$ is the derivative of F w.r.t. the i^{th} argument of F.

Example (Black Scholes model)**.** In the Black-Scholes framwork, the Stock price is modelled by a Smooth Brownian functional sampled at one point $W(T)$.
$F(W(T)) = S(T) = S(0)e^{\left(r-\frac{1}{2}\sigma^2\right)T+\sigma W(T)}$

A direct application of the definition before yields to the Malliavin Derivative

$$
\begin{aligned}
D_t S(T) &= \frac{\partial F(W(T))}{\partial W(T)} \mathbf{1}_{[t,\infty[}(T) \\
&= \sigma S(0) exp \left(\left(r - \frac{1}{2}\sigma^2 \right) T + \sigma W(T) \right) \mathbf{1}_{[t,\infty[}(T) = \sigma S(T) \mathbf{1}_{[t,\infty[}(T).
\end{aligned}
$$

4.5.3 The Skorohod Integral

The *Skorohod integral* is the adjoint operator of the Malliavin Derivative and is also often called *divergence operator*. For the existence of this operator see [37].
The Domain of Malliavin Derivatives is the Banach space $\mathbb{D}_{1,2}$, which is the completion of the set of smooth Brownian Functionals with respect to the Operator norm

$$
||F||_{1,2} = \left(\mathbb{E}|F|^2 + \mathbb{E} \left(\int_0^T |D_t F|^2 dt \right) \right)^{\frac{1}{2}}.
$$

Furthermore the derivative operator D is a closed, linear operator on a dense domain, and therefore one can define its adjoint.

Definition (Skorohod Integral)**.** The Skorohod integral $D^* : \mathcal{D}^{1,2} \to L^2(\Omega)$ has domain

$$
\mathcal{D}^{1,2} \quad := \quad \left\{ u \in L^2(\Omega, \mathcal{F}, P) \otimes H : \mathbb{E}(\langle DF, u \rangle_H) := \mathbb{E} \left(\int_0^\infty D_t F u(t) dt \right) \leq c(u) ||F||_2 \right\}
$$

for a stochastic process u, and any $F \in \mathbb{D}_{1,2}$. It is the adjoint operator of the Malliavin derivative and is characterized by the identity

$$
E \left[\int_0^T (D_t F) u_t dt \right] = E[F D^*(u)]. \tag{4.5.1}
$$

If $u \in \mathcal{D}^{1,2}$, then to stochastic process u is said to be Skorohod integrable.

Corollary. *If u_t is $\mathcal{F}_t$-adapted, then the Skorohod integral coincides with the Itô integral, i.e. $D^*(u) = \int_0^T u_t dW_t$ and it holds*

$$
E \left[\int_0^T (D_t F) u_t dt \right] = E[F \int_0^T u_t dW_t]. \tag{4.5.2}
$$

4.5.4 Integration by parts procedure

The use of Malliavin calculus for calculating derivatives of options relies on formula 4.5.2, which gives you a relationship between the Malliavin derivative D and its adjoint operator D^*. This formula allows us to exchange the stochastic derivative acting on F, by a stochastic integral that does no longer affect F. The following procedure can be applied for calculating the Greeks with Malliavin calculus and can be found in detail in [26].

Let us consider two random variables X, Y and our objective is to calculate $E[H'(X)Y]$. Further we assume that the derivative of the function H is not obtainable. Therefore we want to remove the derivative from H in exchange for a new random variable G

$$E[H'(X)Y] = E[H(X)G].$$

Let us define $Z = H(X)$ and therfore $D_tZ = H'(X)D_tX$ and $YD_tZ = YH'(X)D_tX$. So

$$\int_0^T YD_tZdt = \int_0^T YH'(X)D_tXdt = H'(X)Y\int_0^T D_tXdt$$

and therefore

$$H'(X)Y = \int_0^T \frac{YD_tZ}{\int_0^T D_vXdv}dt.$$

For the objective we therefore have

$$E[H'(X)Y] = E\left[\int_0^T (D_tZ)u_tdt\right]$$

with

$$u_t = \frac{Y}{\int_0^T D_vXdv}.$$

Now we can apply the adjoint operator according to equation 4.5.1, and therefore remove the derivative from f in order to get a new random variable G

$$E[H'(X)Y] = E\left[\underbrace{H(X)}_{Z}D^*\left(\frac{Y}{\int_0^T D_vXdv}\right)\right]$$

$$G \equiv D^*\left(\frac{Y}{\int_0^T D_vXdv}\right).$$

Example (Black Scholes Delta, see [26]).

$$S(T) = S(0)e^{\mu T + \sigma W(T)}$$

$$\Delta = \frac{\partial}{\partial S(0)} E\left[e^{-rT} H(S(T))\right] = e^{-rT} E\left[H'(S(T))\frac{S(T)}{S(0)}\right] = \frac{e^{-rT}}{S(0)} E\left[H'(S(T))S(T)\right]$$

$$\Delta = \frac{e^{-rT}}{S(0)} E\left[H(S(T))D^*\left(\frac{S(T)}{\int_0^T D_v S(T)dv}\right)\right]$$

Recall that$D_u S(T) = \sigma S(T)\mathbf{1}_{[u,\infty[}(T)$ and therefore $\int_0^T D_u S(T)du = \sigma T S(T)$. Applying formula 4.5.2, we get

$$D^*\left(\frac{S(T)}{\int_0^T D_v S(T)dv}\right) = D^*\left(\frac{1}{\sigma T}\right) = \frac{W(T)}{\sigma T}$$

and therefore

$$\Delta = e^{-rT} E\left[H(S(T))\frac{W(T)}{S(0)\sigma T}\right].$$

In [26] it is shown, that if we consider European Style Options and if there is a closed formula for the probability density function of $S(T)$, then the Malliavin calculus reduces to the Score Function method, otherwise the score is replaced by the Skorohod integral. Advantages of the Malliavin calculus are that the payoff function may be discontinuous and the density has not to be known.
The disadvantages are one the one hand that it could be computational very demanding to compute the Skorohod integral, see [12]. On the other hand the Malliavin calculus produces often high variance, [26].

4.6 Measure valued differentiation

4.6.1 Introduction

The Measure valued Differentiation method (MVD) is based on a weak notion of differentiability of probability measures. Let θ be the parameter of interest. We want to estimate

$$\frac{d}{d\theta}\mathbb{E}[H(S_\theta)] = \int H(\omega)\frac{d}{d\theta}\nu_\theta(d\omega).$$

where ν_θ is the distribution of S_θ under θ.

Definition. Let $\mathcal{H}$ be a set of mappings $H : \mathbb{R} \mapsto \mathbb{R}$ which are absolutely integrable with respect to ν_θ for any $\theta \in \Theta \in \mathbb{R}, \mathcal{H} \subset L^1(\nu_\theta, \Theta)$. A function $\theta \mapsto \nu_\theta$ mapping an open subset of $\mathbb{R}$ into the family of all probability measures is called $\mathcal{H}$-differentiable (weakly differentiable w.r.t. $\mathcal{H}$), if a finite signed measure ν'_θ exists, such that

$$\forall H \in \mathcal{H}: \qquad \lim_{h\to 0} \frac{1}{h}\Big(\int H(s)\nu_{\theta+h}(ds) - \int H(s)\nu_\theta(ds)\Big) = \int H(s)\nu'_\theta(ds).$$

Let c_θ be a constant and ν_θ^+ and ν_θ^- two probability measures such that

$$\int H(s)\nu'_\theta(ds) = c_\theta\Big(\int H(s)\nu_\theta^+(ds) - \int H(s)\nu_\theta^-(ds)\Big)$$

for all $H \in \mathcal{H}$, then the triplet $(c_\theta, \nu_\theta^+, \nu_\theta^-)$ is called a *weak derivative triplet* of ν_θ.

The probability measures ν_θ^+ and ν_θ^- can be obtained by decomposing $d\nu_\theta/d\theta$ into a difference between two densities. Such a decomposition can always be found by the Jordan-Hahn decomposition of signed measures, see [41], but any other decomposition may also do the job. A $\mathcal{H}$-derivative refers to an unbiased gradient estimator.

Let us summarize the method: Suppose that S_θ is distributed according to ν_θ and that $H \in \mathcal{H}$. Let us further assume that $\theta \mapsto \nu_\theta$ is weakly differentiable with triplet $(c_\theta, \nu_\theta^+, \nu_\theta^-)$. Then the fundamental equation

$$\frac{\partial}{\partial\theta}\mathbb{E}[H(S_\theta)] = c_\theta[H(S_\theta^+) - H(S_\theta^-)] \tag{4.6.1}$$

holds, provided that the random variables realize the weak derivative, i.e. they satisfy $S_\theta \sim \nu_\theta^+$ and $S_\theta \sim \nu_\theta^-$. In our terminology, we call the random variables S_θ^+ resp. S_θ^- the *positive* resp. *negative realization* of the MVD.

There are numerous of papers, where the theory of MVD has been established, [3, 41, 17, 4]. The interchange of expectation and differentiation is widely justified, because probability densities tend to be smooth functions of their parameters. The condition for unbiasedness - the integrability condition- is a simple growth condition on the payoff function H, see [3]. Further discussions on $\mathcal{H}$ are made in subsection 4.6.4. MVD is independent of the choice of the payoff function H.

If we include the indicator mappings into $\mathcal{H}$, then weak differentiability and strong differentiability coincide. So, MVD is applicable when the score function is and vice versa. The advantage of MVD compared to the SF method is in terms of variance, see [5].

There are also cases where MVD can be applied while SF cannot. This is the instance if we consider a uniform distributed random variable and H out of the set of continuous functions, which is in praxis usually not the case.

Example. Let X_θ be uniformly distributed on $[0,\theta]$ and let $\mathcal{U}_{[0,\theta]}$ denote the uniform distribution on the interval $[0,\theta]$. By $\mathcal{C}[0,\theta]$ denote the set of continuous mappings from $[0,\theta]$ to $\mathbb{R}$. The density of X_θ is

$$f_\theta(x) = \frac{1}{\theta}\mathbf{1}_{[0,\theta]}(x)$$

which is not differentiable with respect to θ, and therefore the Score Function method doesn't apply. For MVD note that for any mapping $H \in \mathcal{C}_{[0,\theta]}$

$$\begin{aligned}
\frac{d}{d\theta}\int H(x)\mathcal{U}_{[0,\theta]}(dx) &= \frac{d}{d\theta}\int H(x)f_\theta(x)dx = \frac{d}{d\theta}\left(\frac{1}{\theta}\int_0^\theta H(x)dx\right) \\
&= -\frac{1}{\theta^2}\int_0^\theta H(x)dx + \frac{1}{\theta}H(\theta) \\
&= \frac{1}{\theta}\left(H(\theta) - \frac{1}{\theta}\int_0^\theta H(x)dx\right) = \frac{1}{\theta}\left(H(\theta) - \int H(x)f_\theta(x)dx\right) \\
&= \frac{1}{\theta}\left(\int H(x)\delta_\theta(dx) - \int H(x)\mathcal{U}_{[0,\theta]}(dx)\right).
\end{aligned}$$

Hence $(1/\theta, \delta_\theta, \mathcal{U}_{[0,\theta]})$ is a weak $\mathcal{C}[0,\theta]$-derivative and we get

$$\frac{d}{d\theta}\mathbb{E}[H(X_\theta)] = \frac{1}{\theta}\mathbb{E}[H(\theta) - H(X_\theta)].$$

4.6.2 Examples for weak derivatives

Example. Poisson distribution
Let X_θ be Poisson(θ) distributed, therefore $\mathbb{P}(X_\theta = k) = \frac{\theta^k}{k!}e^{-k}$. Let $H \in \mathcal{H}$, then

$$\begin{aligned}
\frac{d}{d\theta}\left(\sum_{k=0}^\infty H(k)e^{-\theta}\frac{\theta^k}{k!}\right) &= \sum_{k=0}^\infty\left(-H(k)e^{-\theta}\frac{\theta^k}{k!} + H(k)e^{-\theta}\frac{\theta^{k-1}}{k!}k\right) \\
&= \sum_{k=0}^\infty\left(H(k)e^{-\theta}\frac{\theta^{k-1}}{(k-1)!} - H(k)e^{-\theta}\frac{\theta^k}{k!}\right)
\end{aligned}$$

Hence $(1, \text{Poisson}(\theta)+1, \text{Poisson}(\theta))$ is a weak $\mathcal{H}$-derivative and we have

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(X)] = \mathbb{E}[H(X+1) - H(X)].$$

Example (Normal distribution$X \sim N(\mu_\theta, \sigma^2)$)**.**

Let $\mu_\theta = \theta$. The Normal(θ, σ^2) distribution has density

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right).$$

For the weak derivative we differentiate the densitiy function

$$\frac{d}{d\theta}f(x) = \frac{1}{\sigma\sqrt{2\pi}}\frac{x-\theta}{\sigma^2}exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right).$$

For $\theta = 0$ (for other values of θ we just have to shift) we get

$$\frac{1}{\sigma\sqrt{2\pi}}\frac{x}{\sigma^2}exp\left(-\frac{x^2}{2\sigma^2}\right) = \frac{1}{\sigma\sqrt{2\pi}}\left[\frac{x}{\sigma^2}exp\left(-\frac{x^2}{2\sigma^2}\right)\mathbf{1}_{[x\geq 0]} - \frac{-x}{\sigma^2}exp\left(-\frac{x^2}{2\sigma^2}\right)\mathbf{1}_{[x<0]}\right].$$

This yields for $X \sim N(\mu_\theta, \sigma^2)$ and $H \in \mathcal{H}$

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(X)] = \underbrace{\frac{\mu'_\theta}{\sigma\sqrt{2\pi}}}_{c_\theta}\left[\mathbb{E}_\theta[H(X^+)] - \mathbb{E}_\theta[H(X^-)]\right]$$

with

$$\begin{aligned} X^+ \sim& \mu_\theta + \text{Rayleigh}(\sigma) \\ X^- \sim& \mu_\theta - \text{Rayleigh}(\sigma). \end{aligned}$$

Note that $\text{Rayleigh}(\sigma)$ is $\sqrt{\text{Gamma}(1, 2\sigma^2)}$.

Example ($X \sim N(\mu, \sigma_\theta^2)$). Let $\sigma_\theta = \theta$. If we consider the parameter of interest θ to be in the diffusion, we have to differentiate in the following way:

$$\frac{d}{d\theta}f(x) = -\frac{1}{\theta^2\sqrt{2\pi}}exp\left(-\frac{(x-\mu)^2}{2\theta^2}\right) + \frac{1}{\theta\sqrt{2\pi}}\left(\frac{(x-\mu)^2}{\theta^3}\right)exp\left(-\frac{(x-\mu)^2}{2\theta^2}\right).$$

So for $\mu = 0$ we obtain

$$\begin{aligned} \frac{d}{d\theta}f(x) &= -\frac{1}{\theta^2\sqrt{2\pi}}exp\left(-\frac{x^2}{2\theta^2}\right) + \frac{1}{\theta\sqrt{2\pi}}\frac{x^2}{\theta^3}exp\left(-\frac{x^2}{2\theta^2}\right) \\ &= \frac{1}{\theta}\left[\frac{x^2}{\theta^3\sqrt{2\pi}}exp\left(-\frac{x^2}{2\theta^2}\right) - \frac{1}{\theta\sqrt{2\pi}}exp\left(-\frac{x^2}{2\theta^2}\right)\right]. \end{aligned}$$

This yields for $X \sim N(\mu, \sigma_\theta^2)$ and $H \in \mathcal{H}$

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(X)] = \frac{\sigma'_\theta}{\sigma_\theta}\left[\mathbb{E}_\theta[H(X^+)] - \mathbb{E}_\theta[H(X^-)]\right]$$

with

$$\begin{aligned} X^+ &\sim \mu + \text{ds-Maxwell}(0, \sigma_\theta^2) \\ X^- &\sim \mu + \text{Normal}(0, \sigma_\theta^2) = \text{Normal}(\mu, \sigma_\theta^2) \end{aligned}$$

Example ($X \sim N(\mu_\theta, \sigma_\theta)$).

Now let us assume that the parameter θ occurs in the drift and in the diffusion part, and $\mu'_\theta > 0$ and $\sigma'_\theta > 0$. Then

$$\frac{d}{d\theta}\mathbb{E}_\theta[H(X)] = c_\theta\left[\mathbb{E}_\theta[H(X^+)] - \mathbb{E}_\theta[H(X^-)]\right]$$

where we have with probability p that

$$\begin{aligned} X^+ &\sim \mu_\theta + \text{Rayleigh}(\sigma_\theta) \\ X^- &\sim \mu_\theta - \text{Rayleigh}(\sigma_\theta) \end{aligned}$$

and with probability $1-p$

$$\begin{aligned} X^+ &\sim \mu_\theta + \text{ds-Maxwell}(2\sigma_\theta^2) \\ X^- &\sim \text{Normal}(\mu_\theta, \sigma_\theta^2). \end{aligned}$$

The probability p we therefore calculated by

$$p = \frac{\frac{\mu'_\theta}{\sigma_\theta\sqrt{2\pi}}}{\frac{\mu'_\theta}{\sigma_\theta\sqrt{2\pi}} + \frac{\sigma'_\theta}{\sigma_\theta}} = \frac{\mu'_\theta}{\mu'_\theta + \sigma'_\theta\sqrt{2\pi}}$$

and hence

$$1 - p = \frac{\sigma'_\theta\sqrt{2\pi}}{\mu'_\theta + \sigma'_\theta\sqrt{2\pi}}.$$

For c_θ we have

$$c_\theta = \mu'_\theta \frac{1}{\sigma_\theta\sqrt{2\pi}} + \frac{\sigma'_\theta}{\sigma_\theta} = \frac{\mu'_\theta + \sigma'_\theta\sqrt{2\pi}}{\sigma_\theta\sqrt{2\pi}} \tag{4.6.2}$$

4.6.3 MVD via the characteristic function

Alternatively to the differentiation of measures, one may also differentiate the characteristic functions: Suppose that φ_θ is the characteristic function of ν_θ and suppose further that its derivative w.r.t. θ can be written as

$$\frac{\partial}{\partial\theta}\varphi_\theta(u) = c_\theta(\varphi_\theta^+(u) - \varphi_\theta^-(u))$$

where φ_θ^+ and φ_θ^- are characteristic functions as well. Then we have found the $\mathcal{H}$-weak derivative triplet. Before listing some examples, we review important relationships between some probability distributions.

- If $X \sim \text{Gamma}(1/2, 2\sigma^2)$ (i.e. a multiple of a χ^2 distribution with one degree of freedom), then $\sqrt{X}$ is distributed as the absolute value of a $N(0, \sigma^2)$ distribution.
- If $X \sim \text{Gamma}(1, 2\sigma^2)$, (i.e. a multiple of a χ^2 distribution with two degrees of freedom), then $\sqrt{X}$ is distributed according to a Raleigh(σ^2) distribution, which is a Weibull distribution with exponent 2. The Raleigh(σ^2) distribution has density $\frac{x}{\sigma^2}\exp(-\frac{x^2}{2\sigma^2})\,\mathbf{1}_{\{x\geq 0\}}$ and characteristic function

  $$\varphi(u) = 1 + iu\sigma\sqrt{2\pi}\exp\left(-\frac{u^2\sigma^2}{2}\right)\Phi(iu\sigma),$$

  where Φ is the analytic continuation of the normal Gaussian distribution function.

Distribution	Characteristic function $\varphi(u)$
Exponential(λ)	$\varphi(u) = \frac{\lambda}{\lambda - iu}$
Erlang(λ, n)	$\varphi(u) = \frac{\lambda^n}{(\lambda - iu)^n}$
Normal(μ, σ^2)	$\varphi(u) = exp\left[iu\mu - \frac{\sigma^2}{2}u^2\right]$
ds-Maxwell(σ^2)	$\varphi(u) = \left(1 - \sigma^2 u^2\right) exp\left[-\frac{u^2}{2}\sigma^2\right]$
Gamma(k, θ)	$\varphi(u) = (1 - iu\theta)^{-k}$
Poisson(λ)	$\varphi(u) = exp(\lambda(e^{iu} - 1))$
Raleigh(σ^2)	$\varphi(u) = 1 + iu\sigma\sqrt{2\pi}\exp\left(-\frac{u^2\sigma^2}{2}\right)\Phi(iu\sigma)$

Table 4.1: Distributions and their characteristic functions.

- If $X \sim$ Gamma$(3/2, 2\sigma^2)$, (i.e. a multiple of a χ^2 distribution with three degrees of freedom), then $\sqrt{X}$ is distributed according to a Maxwell(σ^2) distribution. The Maxwell(σ^2) distribution has density $\sqrt{\frac{2}{\pi}}\frac{x^2}{\sigma^3}\exp\left(-\frac{x^2}{2\sigma^2}\right)\mathbf{1}_{\{x\geq 0\}}$ and characteristic function $\varphi(u) = \sqrt{\frac{2}{\pi}}iu\sigma + 2\exp\left(-\frac{u^2\sigma^2}{2}\right)\Phi(iu\sigma)(1 - u^2\sigma^2)$. The symmetrized and shifted version of it is called double sided Maxwell distribution (ds-Maxwell(μ, σ^2)). It has density $\frac{1}{\sigma^3\sqrt{2\pi}}(x - \mu)^2\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ and characteristic function $\varphi(u) = (1 - u^2\sigma^2)\exp\left(iu\mu - \frac{u^2\sigma^2}{2}\right)$.

We are now ready to derive some weak derivative triplets.

- Normal distribution $N(\mu, \sigma^2)$.
  Differentiating the characteristic function of the Normal $N(\mu, \sigma^2)$ distribution w.r.t. μ yields

$$\frac{\partial}{\partial\mu}\exp\left(iu\mu - \frac{\sigma^2}{2}u^2\right) = iu\exp\left(iu\mu - \frac{\sigma^2}{2}u^2\right)$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left\{\underbrace{\exp(iu\mu)\left[1 + iu\sigma\sqrt{2\pi}\exp\left(-\frac{u^2\sigma^2}{2}\right)\Phi(iu\sigma)\right]}_{\mu+\text{Raleigh}(\sigma^2)}\right.$$

$$\left. - \underbrace{\exp(iu\mu)\left[1 - iu\sigma\sqrt{2\pi}\exp\left(-\frac{u^2\sigma^2}{2}\right)\Phi(-iu\sigma)\right]}_{\mu-\text{Raleigh}(\sigma^2)}\right\}.$$

Differentiating the characteristic function of the Normal $N(\mu, \sigma^2)$ distribution w.r.t.

σ yields

$$\frac{\partial}{\partial\sigma}\exp\left[iu\mu-\frac{\sigma^2}{2}u^2\right]=(-u^2\sigma)\exp\left[iu\mu-\frac{\sigma^2}{2}u^2\right]$$

$$=\frac{1}{\sigma}\left(\underbrace{(1-\sigma^2u^2)\exp\left[iu\mu-\frac{\sigma^2}{2}u^2\right]}_{\text{ds-Maxwell}(\mu,\sigma^2)}-\underbrace{\exp\left[iu\mu-\frac{\sigma^2}{2}u^2\right]}_{\text{Normal}(\mu,\sigma^2)}\right).$$

- Poisson(λ) distribution.
  Differentiating the characteristic function of the Poisson(λ) distribution w.r.t. λ yields

$$\begin{aligned}\frac{\partial}{\partial\lambda}\exp[\lambda(e^{iu}-1)] &= (e^{iu}-1)\exp[\lambda(e^{iu}-1)]\\ &= \underbrace{e^{iu}\exp[\lambda(e^{iu}-1)]}_{\text{Poisson}(\lambda)+1}-\underbrace{\exp[\lambda(e^{iu}-1)]}_{\text{Poisson}(\lambda)}\end{aligned}$$

  where $e^{iu}\exp[\lambda(e^{iu}-1]$ is the characteristic function of a Poisson(λ) random variable shifted by 1, since $\mathbb{E}(e^{iu(X+1)})=e^{iu}\exp[\lambda(e^{iu}-1]$ with $X\sim$ Poisson(λ).

- Gamma(a,b) distribution.
  Differentiating the characteristic function of the Gamma (a,b) distribution w.r.t. the scale parameter b gives

$$\frac{\partial}{\partial b}(1-iub)^{-a}=iua(1-iub)^{-a-1}=\frac{a}{b}\left(iub(1-iub)^{-(a+1)}\right)$$

$$=\frac{a}{b}\left(\underbrace{(1-iub)^{-(a+1)}}_{\text{Gamma}(a+1,b)}-\underbrace{(1-iub)^{-a}}_{\text{Gamma}(a,b)}\right)$$

- Exponential distribution. Differentiating the Exponential(λ) with density $\lambda\exp(-\lambda x)$ and characteristic function $\varphi(u)=\lambda/(\lambda-iu)$ w.r.t. λ yields

$$\frac{\partial}{\partial\lambda}\left(\frac{\lambda}{\lambda-iu}\right)=\frac{1}{\lambda-iu}-\frac{\lambda}{(\lambda-iu)^2}=\frac{1}{\lambda}\left(\underbrace{\frac{\lambda}{\lambda-iu}}_{\text{Exponential}(\lambda)}-\underbrace{\frac{\lambda^2}{(\lambda-iu)^2}}_{\text{Erlang}(\lambda,2)}\right)$$

  The Erlang($\lambda,2$) distribution is the same as the Gamma($2,1/\lambda$) distribution.

We summarize the results in Table 4.2 . We can see that one of the realizations is often distributed as the nominal part, which reduces the compuational effort.

Distribution ν_θ (θ varies)	c_θ	Positive Part ν_θ^+	Negative Part ν_θ^-
Poisson(θ)	1	Poisson(θ)+1	Poisson(θ)
Normal(θ, σ^2)	$1/\sigma\sqrt{2\pi}$	θ+Rayleigh(σ)	θ−Rayleigh(σ)
Normal(μ, θ^2)	$1/\theta$	ds-Maxwell(μ, θ^2)	Normal(μ, θ^2)
Gamma(a, θ)	a/θ	Gamma($a+1, \theta$)	Gamma(a, θ)
Exponential(θ)	$1/\theta$	Exponential(θ)	Erlang(θ,2)

Table 4.2: Examples of Weak derivatives

Remark 2. For a given continuous transform $\mathbb{T}$, the image measure (push-forward measure) $\nu^{\mathbb{T}}$ of measure ν is defined as

$$\nu^{\mathbb{T}}(A) = \nu(\mathbb{T}^{-1}(A)).$$

If $\theta \mapsto \nu_\theta$ is differentiable in the measure valued sense with derivative triplet $(c_\theta, \nu_\theta^+, \nu_\theta^-)$, then also $\theta \mapsto \nu_\theta^{\mathbb{T}}$ is differentiable and has triplet $(c_\theta, (\nu_\theta^+)^{\mathbb{T}}, (\nu_\theta^-)^{\mathbb{T}})$. To put this into the level of realizations, if the triplet $(c_\theta, X_\theta^+, X_\theta^-)$ realizes $(c_\theta, \nu_\theta^+, \nu_\theta^-)$, then $(c_\theta, \mathbb{T}(X_\theta^+), \mathbb{T}(X_\theta^-))$ realizes $(c_\theta, (\nu_\theta^+)^{\mathbb{T}}, (\nu_\theta^-)^{\mathbb{T}})$.

This fact is especially important when one considers the transformation of a Lévy process to an exponential Lévy process using the transform $\mathbb{T}(w) = S_0 \exp(w)$. If the two random variables X^+ resp. X^- realize the two parts of the MVD of ν_θ, then $S_0 \exp(X^+)$ resp. $S_0 \exp(X^-)$ realize the two parts of the MVD of $\nu_\theta^{\mathbb{T}}$, i.e. of the exponential model.

4.6.4 The function set $\mathcal{H}$

The concept of weak derivative states that $\theta \mapsto \nu_\theta$ is weakly differentiable with respect to $\mathcal{H}$, if a finite signed measure ν'_θ exists, such that

$$\frac{\partial}{\partial\theta} \int H(s)\nu_\theta(ds) = \int H(s)\nu'_\theta(ds)$$

for $H \in \mathcal{H}$. In contrast to strong differentiability, the weak concept is dependent on the two partners H and ν_θ.

One way to impose some structure on the set $\mathcal{H}$ is explained in [3]. Defining a supremums-norm

$$||H||_v = \sup_{s\in\mathcal{S}} \frac{|H(s)|}{v(s)}$$

with v a measurable mapping $v : S \to \mathbb{R}$ with $\inf_{s\in S} v(s) \geq 1$. For a (signed) measure μ, the asscociated measure norm is

$$||\mu||_v = \sup_{||H||_v \leq 1} |\mu H|.$$

Let v_p be a polynomial of degree p and $\mathcal{H}$ a set of measurable mappings. Then $(\mathcal{H}, p)$ is the set of measurable mappings that are bounded by a polynomial of degree p

$$(\mathcal{H}, p) \triangleq \left\{H \in \mathcal{H} : ||H||_{v_p} < \infty\right\}.$$

Let ν_θ be $(\mathcal{H}, p)$-differentiable at θ with $(\mathcal{H}, p)$ Banach. In any neighborhood $U = [\theta - \Delta, \theta + \Delta]$ it then holds that

$$\forall |h| \leq \Delta : \quad \int H(s)|\nu_{\theta+h} - \nu_\theta| ds \leq ||\nu_{\theta+h} - \nu_\theta||_{v_p} v(s)$$

and $||\nu_{\theta+h} - \nu_\theta||_{v_p} \leq |h|M$ for a finite constant M, for all measurable H such that $||H||_{v_p} < \infty$. So a classical example for $(\mathcal{H}, p)$ would be $(\mathcal{B}(s), p)$ which is the set of all measurable mappings bounded by a polynomial of degree p. As already mentioned: If we include the indicator mappings into $\mathcal{H}$, then weak differentiability and strong differentiability coincide. So, MVD is applicable when the score function is and vice versa. This is no problem as the real advantage of MVD over the score function is in terms of variance (and WNV).

In this work we deal with concrete models. For our purposes, a practical way to define the function set $\mathcal{H}$ proceeds as follows. Let $\nu_\theta : \theta \in \Theta \subset \mathbb{R}$ be a family of probability measures, say on $\mathbb{R}^m$ and let $\mathcal{C}_b$ the family of bounded continuous functions on $\mathbb{R}^m$. Introduce the notation

$$\langle H, \nu_\theta \rangle$$

for

$$\int H(u)\, d\nu_\theta(u).$$

We say that $\theta \mapsto \nu_\theta$ is weakly differentiable at θ_0, if there is a signed measure ν'_{θ_0} with $|\nu'_{\theta_0}|(\mathbb{R}) < \infty$ such that for all

$$\frac{\partial}{\partial \theta} \langle H, \nu_\theta \rangle \Big|_{\theta=\theta_0} = \langle H, \nu'_{\theta_0} \rangle. \tag{4.6.3}$$

$H \in \mathcal{C}_b$. This ensures the uniquenenss of the weak derivative[3]. Now, let $\mathcal{H}$ be a family of functions with $\mathcal{C}_b \subseteq \mathcal{H}$. One may define a notion like $\mathcal{H}$-differentiability, if (4.6.3) holds true for all $H \in \mathcal{H}$. But this is not necessary. Notice that the derivative element ν'_θ does not depend on $\mathcal{H}$. In fact, if $\mathcal{H}$-differentiability holds, then also $\mathcal{C}_b$ differentiability holds and the derivatives are the same.

Thus we define the differentiability only with respect to the smallest function family, which is $\mathcal{C}_b$ (as a property of the family ν_θ) and ask then whether the derivative is such that also larger families than $\mathcal{C}_b$ qualify for (4.6.3).

There is no problem to extend the notion to all functions which are bounded and $|\nu'_\theta|$ a.e. continuous. Here is an example, why one cannot get rid of this condition: Let $\nu_\theta = \theta\delta_\theta + (1-\theta)\delta_1$. Then the weak derivative at $\theta = 0$ is $\delta_0 - \delta_1$. However, (4.6.3) does not hold for functions which are discontinuous at 0.

The extension to unbounded H is more involved: Lets say that the derivatives ν'_θ have *locally uniform p-th moments at* θ_0, if there is an open neighborhood Θ_0 of θ_0 such that

[3]Notice however that the decomoposition into the difference of two probability measures is not unique

for every $\epsilon > 0$ there is a K_ϵ, such that

$$\sup_{\theta \in \Theta_0} \langle |u|^p \mathbf{1}_{|u| \geq K_\epsilon}, |\nu'_\theta| \rangle \leq \epsilon. \tag{4.6.4}$$

If H is some continuous function which is bounded by $C(1 + |u|^p)$, then under the condition (4.6.4) the relation (4.6.3) holds for H. Here is the proof: Write H as $H(u) = H_1(u) + H_2(u)$ with H_1 is bounded and continuous and $|H_2(u)| \leq C|u|^p \mathbf{1}_{|u| \geq K_\epsilon}$. Then $\theta \mapsto \langle H_1, \nu_\theta \rangle$ is differentiable. As to the second part, notice that

$$\begin{aligned} & \langle H_2, \nu_{\theta_0+h} - \nu_{\theta_0} - h \cdot \nu'_{\theta_0} \rangle = \langle H_2, \int_{\theta_0}^{\theta_0+h} (\nu'_s - \nu'_{\theta_0})\, ds \rangle \\ \leq \;& \langle C|u|^p \mathbf{1}_{|u| \geq K_\epsilon}, \int_{\theta_0}^{\theta_0+h} (|\nu_s|' + |\nu'_{\theta_0}|)\, ds \rangle \\ \leq \;& 2\epsilon \cdot h \end{aligned}$$

Since ϵ is arbitrary, the result follows.

The Lemma below gives a sufficient condition, which guarantees the condition (4.6.4): Let ϕ_θ be the family of characteristic functions pertaining to the family ν_θ. Then $\theta \mapsto \nu_\theta$ is weakly differentiable if and only if $\phi'_\theta(u) = \frac{\partial}{\partial \theta} \phi_\theta(u)$ exists, such that ϕ'_θ is the difference of two positive definite functions (in the sense of Bochner):

$$\phi'_\theta = c_\theta(\dot{\phi}_\theta - \ddot{\phi}_\theta).$$

Here the normalization is such that $\dot{\phi}_\theta(0) = \ddot{\phi}_\theta(0) = 1$.

Let $\dot{\nu}_\theta$ resp. $\ddot{\nu}_\theta$ be the probability measures pertaining to $\dot{\phi}_\theta$ and $\ddot{\phi}_\theta$, respectively. Notice that $\int \exp(iut) d|\nu'_\theta|(u) = c_\theta(\dot{\phi}_\theta(t) + \ddot{\phi}_\theta(t))$. We require the weak differentiability in a neighborhood Θ_0 of θ_0.

Lemma. Let $q = p + 1$ if p is odd and $q = p + 2$ if p is even. Suppose that

- (i) $(\theta, u) \mapsto \phi_\theta(u)$ is q times differentiable w.r.t. u and once differentiable w.r.t. θ such that for a neighborhood Θ_0 of θ_0

$$\sup_{\theta \in \Theta_0} |\frac{\partial^q}{\partial u^q} [\dot{\phi}_\theta(u) + \ddot{\phi}_\theta(u)]| \Big|_{u=0} < \infty$$

- (ii) $\sup_{\theta \in \Theta_0} c_\theta < \infty$.

Then (4.6.4) is fulfilled.

Proof. Condition (i) implies that the q-th moments of $\dot{\nu}_\theta$ and $\ddot{\nu}_\theta$ are uniformly bounded. Hence, if $q = p + 1$

$$\sup_{\theta \in \Theta_0} \langle |u|^p \mathbf{1}_{|u| \geq K_\epsilon}, \dot{\nu}_\theta(u) \rangle \leq \frac{1}{K_\epsilon} \sup_{\theta \in \Theta_0} \int u^q \, d\dot{\nu}_\theta(u) \leq \epsilon$$

if K_ϵ is large enough. The same argument for $\ddot{\nu}_\theta$ and condition (ii) implies the result. The case $q = p + 2$ is analogous.

Fact 3. *Suppose $X \sim \nu_\theta$ has characteristic function $\varphi_\theta(u)$.*

If we assume that $(\theta, u) \mapsto \varphi_\theta(u)$ is jointly differentiable in both variables, but at least $p+1$-times w.r.t. u, and if (4.6.4) is fulfilled, then ν_θ is weakly differentiable w.r.t. $\mathcal{H}_p$.

$\mathcal{H}_p$ is the space of all ν'_θ-a.e. continuous functions $\forall \theta \in \Theta_0$, which do not grow faster than $1 + |x|^p$.

This is not a real restriction on the payoff functions H. Whereas for example the Vanilla Call or Put Option, the Lookback Option and the Asian Option do not grow faster than linear (i.e. $p = 1$) and are continuous, the Digital Option is bounded by a constant and its points of discontinuity have in all relevant applications mass 0.

4.6.5 The Measure Valued Derivative for Markov processes and its estimation

As was already said, we consider only discrete time homogeneous Markov processes with infinitely divisible increment distribution in this work. Suppose that Δ is the time increment. With a slight abuse of notation, we write now

$$S_\theta(i) \qquad \text{for} \qquad S_\theta(i\Delta)$$

bearing in mind that the maturity time T corresponds to the step n with $T = n \cdot \Delta$.

To the process $S_\theta(i), i = 0, \ldots, n$ we associate the transition operator

$$\mathbb{P}_\theta(w, A) = P\{S_\theta(i+1) \in A | S_\theta(i) = w),$$

the starting distribution[4] γ and the payoff function H.

Introduce the following notations:

$\gamma\mathbb{P}_\theta$	for the measure $(\gamma\mathbb{P}_\theta)(A) = \int \mathbb{P}_\theta(w, A)\, d\gamma(w)$,
$\mathbb{P}_\theta H$	for the function $(\mathbb{P}_\theta H)(u) = \int H(w)\mathbb{P}_\theta(u, dw)$,
$\mathbb{P}^2_\theta(w, A)$	for the two-step transition $\mathbb{P}^2_\theta(w, A) = \int \mathbb{P}_\theta(v, A)\mathbb{P}_\theta(w, dv)$,
$\mathbb{P}^n_\theta(w, A)$	for the n-step transition as the concatenation of $n-1$ two-step transitions.

Using this notation, we write for the expected payoff at maturity time $T = n \cdot \Delta$

$$\mathbb{E}\left[H(S_\theta(n))\right] = \gamma\, \mathbb{P}^n_\theta\, H.$$

Definition. The Markov transition $\mathbb{P}_\theta(\cdot, \cdot)$ is called weakly $\mathcal{H}$-differentiable, $\mathcal{C}_b \subseteq \mathcal{H}$, if there is a signed transition $\mathbb{P}'_\theta$ such that for all functions $H \in \mathcal{H}$ and every point mass δ_w (i.e. the probability distribution concentrated on the point w)

$$\frac{1}{s}|\delta_w \mathbb{P}_{\theta+s}H - \delta_w \mathbb{P}_\theta H - s \cdot \delta_w \mathbb{P}'_\theta H| \to 0$$

as $s \to 0$.

[4] While γ is in most cases just the point mass δ_{S_0} at S_0, we allow here some slight generalization

The finite signed transition $\mathbb{P}'_\theta$ may be decomposed as

$$\mathbb{P}'_\theta(w, A) = c_\theta(w)[\mathbb{P}^+_\theta(w, A) - \mathbb{P}^-_\theta(w, A)], \tag{4.6.5}$$

where $\mathbb{P}^+_\theta$ and $\mathbb{P}^-_\theta$ are regular Markov transitions.

Remark 4. It is assumed that the Markov transition $\mathbb{P}_\theta(\cdot, \cdot)$ is regular, i.e. $w \mapsto \mathbb{P}_\theta(w, A)$ is measurable $\forall A \in \mathcal{F}$ and the mapping $A \mapsto \mathbb{P}_\theta(w, A)$ is a probability measure $\forall w$. Notice that

$$w \mapsto \mathbb{P}_\theta(w, A) \text{ is measurable } \forall A$$

implies that

$$w \mapsto \int H(y)\mathbb{P}_\theta(w, dy) \text{ is measurable for continuous } H$$

implies that

$$w \mapsto \int H(y)\mathbb{P}'_\theta(w, dy) \text{ is measurable}$$

implies that

$$w \mapsto \mathbb{P}'_\theta(w, A) \text{ is measurable } \forall A \in \text{Baire } \sigma\text{-algebra}^5.$$

Remark 5. For exponentials of processes with independent increments, the situation simplifies considerably. Since $S_\theta(t + \Delta) = S_\theta(t) \cdot \exp(X_\theta(t + \Delta) - X_\theta(t))$, by weakly differentiating the increment distribution $X_\theta(t+\Delta)-X_\theta(t)$, giving a positive realization, a negative realization and a constant c_θ, these can be transformed by the exponential transform (see Remark 2) and this triplet does not depend on the previous state $S_\theta(t)$. Thus due to the independent increment property of Lévy processes, the measure valued differentiation of increments can be done very efficiently.

Recall now the Leibnitz rule for the derivation of a power of operators

$$(\mathbb{P}^n_\theta)' = \sum_{i=1}^{n} \mathbb{P}^{i-1}_\theta \mathbb{P}'_\theta \mathbb{P}^{n-i}_\theta. \tag{4.6.6}$$

Using the decomposition (4.6.5) this can be written in terms of MVD

$$\frac{\partial}{\partial\theta}\mathbb{E}\left[H(S_\theta(n))\right]] = \gamma \left[\sum_{i=1}^{n} \mathbb{P}^{i-1}_\theta c_\theta \left(\mathbb{P}^+_\theta - \mathbb{P}^-_\theta\right) \mathbb{P}^{n-i}_\theta\right] H \tag{4.6.7}$$

see [3].

Formula (4.6.7) can be used in different ways to construct unbiased estimates of $\frac{\partial}{\partial\theta}\gamma\,\mathbb{P}^n_\theta\,H$ as we introduce them below: the exact estimate MVDe, the randomized estimate MVDr and the compromise estimate MVDc(k).

[5] The Baire σ-algebra is generated by the level sets of all continuous functions

4.6.6 Algorithms MVD

The following section introduce the algorithm of Measure valued differentiation. We become acquainted with three different forms of the algorithm, which differs in the end in the efficiency of the estimator (computational time and variance).

4.6.6.1 The exact estimator (MVDe)

he exact method uses estimates for every summand in (4.6.7).

1. Sample $S_\theta(0)$ from the starting distribution γ.
2. Sample n steps with transition $\mathbb{P}_\theta$, giving $S_\theta(1), \ldots, S_\theta(n)$.
3. For all $i = 1, ..., n$ sample one transition step from $S_\theta(i-1)$ with transition $\mathbb{P}_\theta^+$ and one with transition $\mathbb{P}_\theta^-$, giving $S_{\theta,i-1}^+(i)$ resp.$S_{\theta,i-1}^-(i)$. Store the values $c_\theta(i)$.
4. Continue these processes $S_{\theta,i-1}^+(l)$ resp.$S_{\theta,i-1}^-(l)$ for $l = i+1, ..., n$ using transition $\mathbb{P}_\theta$ and a coupling technique (see below).
5. The unbiassed estimate is
$$\sum_{i=1}^{n} c_\theta(i)\left[H(S_{\theta,i-1}^+(n)) - H(S_{\theta,i-1}^-(n))\right] \tag{4.6.8}$$
6. The final estimate is the arithmetic mean of N independent replications of estimate (4.6.8).

The processes $(S_{\theta,i-1}^+(\cdot), S_{\theta,i-1}(\cdot))$ are called *positive phantoms*, resp. *negative phantoms*, together they are the *phantom pairs*. The exact method uses n phantom pairs, which results in a high computational effort.

4.6.6.2 The randomized estimator (MVDr)

One may avoid to evaluate large sums by using the identity
$$\sum_{i=1}^{n} a_i = \mathbb{E}[n a_\tau]$$
for a random τ, which is uniformly distributed on the integers $1, \ldots, n$. In our context, the following identity is used
$$\frac{\partial}{\partial\theta}\mathbb{E}\left[H(S_\theta(n))\right]] = n\mathbb{E}\left[\gamma\left[\left(\mathbb{P}_\theta^{\tau-1} c_\theta\left(\mathbb{P}_\theta^+ - \mathbb{P}_\theta^-\right)\mathbb{P}_\theta^{n-\tau}\right)\right] H\right] \tag{4.6.9}$$
where the expectation is over the random τ. The randomized method simulates only one single phantom pair, which is started at a random point in time.

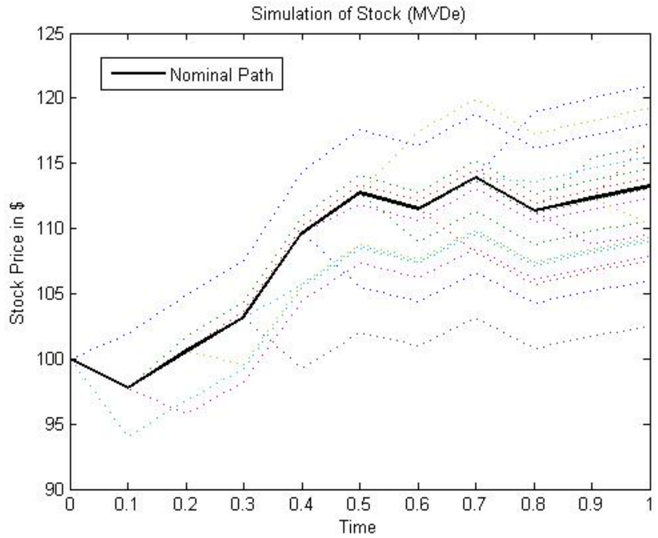


Figure 4.6.1: MVDe, $n = 10$, $t = 1$.

1. Sample a random uniform time τ in $\{1, \ldots, n\}$.
2. Sample $S_\theta(0)$ from the starting distribution γ.
3. Sample $\tau - 1$ steps with transition $\mathbb{P}_\theta$, giving $S_\theta(1), \ldots, S_\theta(\tau - 1)$.
4. Sample one transition step from $S_\theta(\tau-1)$ with transition $\mathbb{P}_\theta^+$ and one with transition $\mathbb{P}_\theta^-$, giving $S_\theta^+(\tau)$ resp. $S_\theta^-(\tau)$. Store c_θ.
5. Continue the processes $S_\theta^+(i)$ resp. $S_\theta^-(i)$, for $i = \tau + 1, \ldots, n$ using transition $\mathbb{P}_\theta$ and a coupling technique.
6. The unbiased estimate is
$$n \cdot c_\theta[H(S_\theta^+(n)) - H(S_\theta^-(n))]. \tag{4.6.10}$$
7. The final estimate is the arithmetic mean of N independent replications of estimate (4.6.10).

MVDr tends to have a higher variance, because of introducing an additional random variable τ. In comparison to the randomized estimator, the MVDe algorithm has smaller variance but because of summing up n elements, the computational effort is much higher than those of the MVDr algorithm. The MVDc(k) algorithm is a compromise between the randomized estimator and the exact estimator.

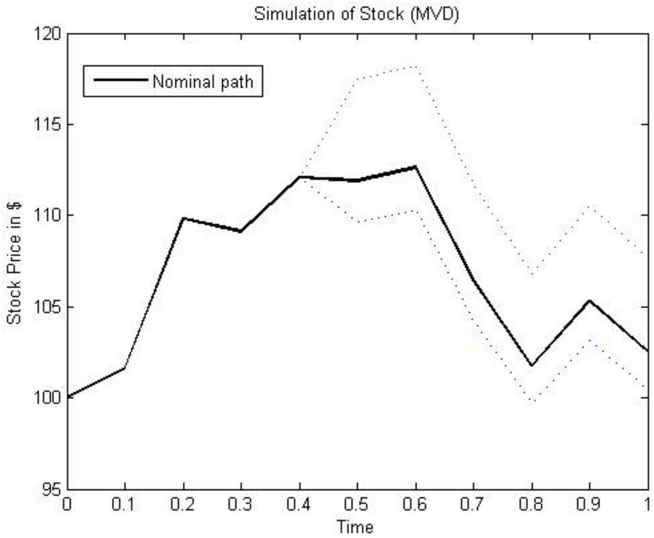


Figure 4.6.2: MVD (random point), $n = 10$, $t = 1$, $\tau = 4$.

4.6.6.3 The compromise estimate (MVDc(k))

This algorithm is a compromise between MVDe and MVDr. It generates exactly k phantom pairs, for $1 < k < n$. Suppose that $n = k \cdot m$.

1. Sample k random uniform times $\tau_1, ..., \tau_k$ in $(1, ..., m)$,...,$(m(i-1)+1, ..., mi)$,..., $((k-1)m+1, ..., km = n)$
2. Sample $S_\theta(0)$ from the starting distribution γ.
3. Sample n steps with transition $\mathbb{P}_\theta$, giving $S_\theta(0), \dots, S_\theta(n)$
4. For all $i = \tau_1, ..., \tau_k$ sample one transition step from $S_\theta(i-1)$ with transition $\mathbb{P}_\theta^+$ and one with transition $\mathbb{P}_\theta^-$, giving $S^+_{\theta,i-1}(i)$ resp.$S^-_{\theta,i-1}(i)$. Store $c_\theta(i)$.
5. Continue these processes $S^+_{\theta,i-1}(l)$ resp.$S^-_{\theta,i-1}(l)$, $l = i+1, ..., n$ using transition $\mathbb{P}_\theta$ and a coupling technique.
6. The unbiased estimate is
$$m \cdot \sum_{i=1}^{k} c_\theta(i) \left[H(S^+_{\theta,i-1}(n)) - H(S^-_{\theta,i-1}(n)) \right]. \tag{4.6.11}$$
7. The final estimate is the arithmetic mean of N independent replications of estimate 4.6.11.

Notice that the MVDc(k) algorithms includes the extreme cases MVDr equal MVDc(1) and MVDe equals MVDc(n).

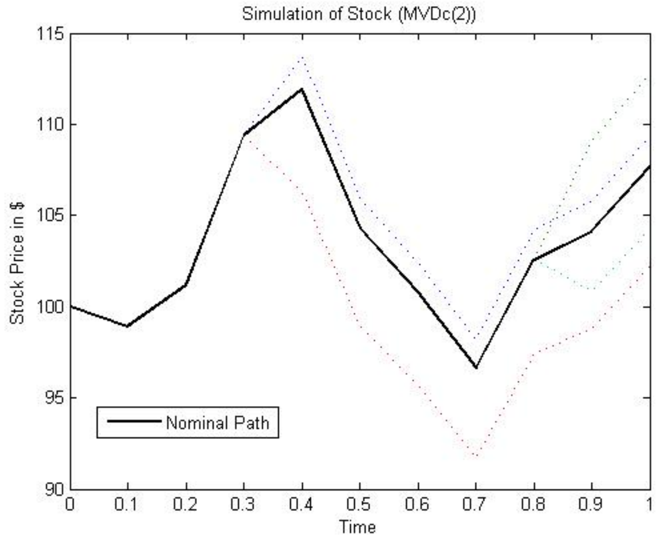


Figure 4.6.3: MVDc(2), $n = 10$, $t = 1$, $k = 2$, $m = 5$.

The MVDc(k) algorithm is a compromise between the randomized estimator and the exact estimator.

4.6.7 Variance reduction via coupling

The variance of a MVD estimator of the form $c \cdot (H(S^+) - H(S^-))$ is given by

$$c^2 \cdot \Big(Var\left[H(S^+)\right] + Var\left[H(S^-)\right] - 2Cov\left[H(S^+), H(S^-)\right] \Big).$$

An appropriate choice of the positive and negative phantoms may cause positive correlation and therefore lead to a reduced variance of the estimator.

Example. Sensitivity estimation w.r.t. μ in the Gaussian case:
Consider a normal distributed random variable $X \sim N(\mu, \sigma^2)$. When calculating the sensitivity w.r.t. μ, the variable splits up into $X^+ \sim \mu + \text{Rayleigh}(\sigma)$ and $X^- \sim \mu - \text{Rayleigh}(\sigma)$ according to the weak derivative triplet, see Table 4.2. For coupling we have just to use the same Rayleigh(σ) for the positive and negative part, see [3, 5].

Example. Sensitivity w.r.t. σ in the Gaussian case:
Consider $X \sim N(\mu, \sigma^2)$ and we are interested in the sensitivity w.r.t. the volatility σ. According to Table 4.2 we have that $X^+ \sim \mu + \text{ds-Maxwell}(0, \sigma^2)$ and $X^- \sim \mu + N(0, \sigma^2)$. According to [5], we use following coupling method: First simulate $V \sim \text{ds-Maxwell}(0, \sigma^2)$ and set $X^+ = \mu + V$. Then sample an independent uniform random variable U on $(0, 1)$ and set $X^- = \mu + UV$. Then X^- is normally distributed with parameters μ and σ and is positively correlated with X^+.

Example. Sensitivity w.r.t. λ in the Poisson case:
As it was already shown, a pair realizing the weak derivative in the Poisson case can be chosen as $X^- \sim \text{Poisson}(\lambda)$ and $X^+ = X^- + 1$.

Example. Sensitivity w.r.t. b in the Gamma case:
If X^- is generated according to Gamma(a, b), then $X^+ \sim \text{Gamma}(a+1, b)$ may be generated as $X^+ = X^- + V$, where $V \sim \text{Exponential}(1/b)$.

4.6.8 MVD for path dependent payoff functions

In this section, we demonstrate how the MVD method can be modified to allow the sensitivity estimation of Lookback and Asian options. These options are *path dependent*, i.e. the payoff depends on the whole history of the underlying and not just at its value at maturity.

4.6.8.1 Estimation via joint transition densities

In the following we use the notation

$$\mathbb{E}\left[H(S_\theta(T))\right] = \mathbb{E}\left[\overline{H}(X_\theta(T))\right]$$

with

$$\overline{H}(w) = H\left(S(0) \cdot e^w\right) \tag{4.6.12}$$

Homogeneous case

Suppose now we want to differentiate an expectation of the form

$$\mathbb{E}_\theta[H(S(1), \ldots, S(T))] = \mathbb{E}_\theta[\overline{H}(X(1), \ldots, X(T))],$$

where X_t is a homogeneous Markov process on a measure space (R, ν) with transition density $f_\theta(u|w)$ and starting value u_0, i.e.

$$P_\theta(X(t) \in A | X(t-1) = w) = \int_A f_\theta(u|w)\, d\nu(u) \qquad a.s.$$

The parameter θ varies in an open set Θ.
For a measurable function $(u, w) \mapsto g(u|w)$: $R \times R \to \mathbb{R}$ introduce

$$\|g(\cdot|\cdot)\|_{p,\infty} = \text{esssup}_w \left[\int |g(u|w)|^p\, d\nu(u)\right]^{1/p}. \tag{4.6.13}$$

Notice that

$$|\int \cdots \int \prod_{t=1}^{T} v(u_t) \cdot g_t(u_t|u_{t-1})\, d\nu(u_1) \ldots d\nu(u_T)| \le (\|v\|_q)^T \cdot \prod_{t=1}^{T} \|g_t(\cdot|\cdot)\|_{p,\infty} \tag{4.6.14}$$

Assumptions.

- (i) The transition densities satisfy
$$\sup_{\theta\in\Theta} \|f_\theta(\cdot|\cdot)\|_{p,\infty} < \infty$$
- (ii) The transition densities are differentiable in the following sense: There are functions $\dot{f}_\theta(\cdot|\cdot)$ such that
$$||\dot{f}_\theta(.|.)||_{p,\infty} < \infty$$
$$\|\frac{1}{h}[f_{\theta+h}(\cdot|\cdot) - f_\theta(\cdot|\cdot)] - \dot{f}_\theta(\cdot|\cdot)\|_{p,\infty} \to 0$$
for $h \to 0$.
- (iii) There is a function $v \in L_q(d\nu)$ $(1/p + 1/q = 1)$, such that
$$|\overline{H}(u_1, \dots, u_T)| \le C \prod_{t=1}^{T} v(u_t).$$

Lemma. Under the given assumptions,
$$\frac{\partial}{\partial\theta} \int \dots \int \overline{H}(u_1, \dots, u_T) \prod_{t=1}^{T} f_\theta(u_t|u_{t-1})\, d\nu(u_1) \dots d\nu(u_T) \tag{4.6.15}$$
$$= \sum_{j=1}^{T} \int \dots \int \overline{H}(u_1, \dots, u_T) \prod_{t=1}^{j-1} f_\theta(u_t|u_{t-1}) \cdot \dot{f}_\theta(u_j|u_{j-1}) \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1})\, d\nu(u_1) \dots d\nu(u_T).$$

Proof. Notice that
$$\frac{1}{h} \int \dots \int \overline{H}(u_1, \dots, u_T) \left[\prod_{t=1}^{T} f_{\theta+h}(u_t|u_{t-1}) - \prod_{t=1}^{T} f_\theta(u_t|u_{t-1})\right] d\nu(u_1) \dots d\nu(u_T)$$
$$= \sum_{j=1}^{T} \int \dots \int \overline{H}(u_1, \dots, u_T) \prod_{t=1}^{j-1} f_{\theta+h}(u_t|u_{t-1}) \cdot \frac{1}{h}[f_{\theta+h}(u_j|u_{j-1}) - f_\theta(u_j|u_{j-1})] \times$$
$$\times \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1})\, d\nu(u_1) \dots d\nu(u_T).$$

The difference of this expression and the limiting expression (4.6.15) is given by $\sum_{j=1}^{T} I_j(h)$ with
$$\begin{aligned} I_j(h) &= \int \dots \int \overline{H}(u_1, \dots, u_T) \Bigg\{ \prod_{t=1}^{j-1} f_{\theta+h}(u_t|u_{t-1}) \cdot \frac{1}{h}[f_{\theta+h}(u_j|u_{j-1}) - f_\theta(u_j|u_{j-1})] \times \\ &\quad \times \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1}) - \prod_{t=1}^{j-1} f_\theta(u_t|u_{t-1} \cdot \dot{f}_\theta(u_j|u_{j-1}) \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1}) \Bigg\}\, d\nu(u_1) \dots d\nu(u_T) \\ &= \sum_{k=1}^{j} I_{j,k} \end{aligned}$$

with

$$\begin{aligned}
I_{j,k}(h) &= \int \cdots \int \overline{H}(u_1, \ldots, u_T) \prod_{t=1}^{k-1} f_{\theta+h}(u_t|u_{t-1}) \cdot [f_{\theta+h}(u_k|u_{k-1}) - f_\theta(u_k|u_{k-1})] \times \\
&\quad \times \prod_{t=k+1}^{j-1} f_\theta(u_t|u_{t-1}) \cdot \dot{f}_\theta(u_j|u_{j-1}) \cdot \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1}) \, d\nu(u_1) \ldots d\nu(u_T)
\end{aligned}$$

for $k < j$, and

$$\begin{aligned}
I_{j,j}(h) &= \int \cdots \int \overline{H}(u_1, \ldots, u_T) \prod_{t=1}^{j-1} f_{\theta+h}(u_t|u_{t-1}) \times \\
&\times \left[\frac{1}{h}(f_{\theta+h}(u_j|u_{j-1}) - f_\theta(u_j|u_{j-1})) - \dot{f}_\theta(u_j|u_{j-1})\right] \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1}) \, d\nu(u_1) \ldots d\nu(u_T).
\end{aligned}$$

Notice that

$$\|f_{\theta+h} - f_\theta\|_{p,\infty} \leq \|f_{\theta+h} - f_\theta - h \cdot \dot{f}_\theta\|_{p,\infty} + |h| \|\dot{f}_\theta\|_{p,\infty} = o(h) + O(h) = O(h). \quad (4.6.16)$$

Now, by (4.6.14), for $k < j$

$$\begin{aligned}
|I_{j,k}(h)| &\leq \int \cdots \int C \prod_{t=1}^{T} v(u_t) \cdot \prod_{t=1}^{k-1} f_{\theta+h}(u_t|u_{t-1}) \cdot |f_{\theta+h}(u_k|u_{k-1}) - f_\theta(u_k|u_{k-1})| \times \\
&\times \prod_{t=k+1}^{j-1} f_\theta(u_t|u_{t-1}) \cdot |\dot{f}_\theta(u_j|u_{j-1})| \cdot \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1}) \, d\nu(u_1) \ldots d\nu(u_T) \\
&\leq (\|v\|_q)^T \cdot (\|f_{\theta+h}(\cdot|\cdot)\|_{p,\infty})^{k-1} \cdot (\|f_\theta(\cdot|\cdot)\|_{p,\infty})^{T-(k+1)} \cdot \|f_{\theta+h} - f_\theta\|_{p,\infty} \|\dot{f}_\theta\|_{p,\infty} \\
&= O(h)
\end{aligned}$$

and for $k = j$

$$\begin{aligned}
|I_{j,j}(h)| &\leq \int \cdots \int C \prod_{t=1}^{T} v(u_t) \cdot \prod_{t=1}^{j-1} f_{\theta+h}(u_t|u_{t-1}) \times \\
&\times \left|\frac{1}{h}(f_{\theta+h}(u_j|u_{j-1}) - f_\theta(u_j|u_{j-1}) - \dot{f}_\theta(u_j|u_{j-1})\right| \prod_{t=j+1}^{T} f_\theta(u_t|u_{t-1}) \, d\nu(u_1) \ldots d\nu(u_T) \\
&\leq (\|v\|_q)^T \cdot (\|f_{\theta+h}(\cdot|\cdot)\|_{p,\infty})^{j-1} \cdot (\|f_\theta(\cdot|\cdot)\|_{p,\infty})^{T-j} \cdot \|\frac{1}{h}(f_{\theta+h} - f_\theta) - \dot{f}_\theta\|_{p,\infty}\| \\
&= o(h).
\end{aligned}$$

Putting all pieces together, we have shown that $I_j(h) \to 0$ as $h \to 0$ for all j, i.e. the Lemma is shown.

Suppose that the derivative has a representation

$$\dot{f}_\theta(u|w) = c_\theta[\dot{g}_\theta(u|w) - \ddot{g}_\theta(u|w)]$$

then the following method leads to an unbiased estimate for

$$\frac{\partial}{\partial\theta}\mathbb{E}_\theta[\overline{H}(X(1), \ldots, X(T))].$$

- Sample a stopping time τ uniformly in $\{1, \ldots, T\}$.
- Simulate the process $X(t)$ under θ until time $\tau - 1$. For the ease of notation set $\dot{X}(t) = \ddot{X}(t) = X(t)$ for $t < \tau$.
- Make one step from $X(\tau)$ with transition density $\dot{g}_\theta$ to give $\dot{X}(\tau+1)$ and continue then with f_θ giving $\dot{X}(\tau+k)$.
- Make one step from $X(\tau)$ with transition density $\ddot{g}_\theta$ to give $\ddot{X}(\tau+1)$ and continue then with f_θ giving $\ddot{X}(\tau+k)$.
- The unbiased estimate is

$$T \cdot c_\theta \left(\overline{H}(\dot{X}(1), \ldots, \dot{X}(T)) - \overline{H}(\ddot{X}(1), \ldots, \ddot{X}(T))\right).$$

Inhomogeneous case

For the inhomogeneous case, we have to replace the density $f_\theta(u|w)$ by $f_\theta^{(t)}(u|w)$ for each individual time t. Hence

$$P_\theta(X(t) \in A | X(t-1) = w) = \int_A f_\theta^{(t)}(u|w)\, d\nu(u) \qquad a.s.$$

The Lemma then reads

Lemma.

$$\frac{\partial}{\partial\theta}\int \ldots \int \overline{H}(u_1, \ldots, u_T) \prod_{t=1}^{T} f_\theta^{(t)}(u_t|u_{t-1})\, d\nu(u_1) \ldots d\nu(u_T) \tag{4.6.17}$$

$$= \sum_{j=1}^{T} \int \ldots \int \overline{H}(u_1, \ldots, u_T) \prod_{t=1}^{j-1} f_\theta^{(t)}(u_t|u_{t-1}) \cdot \dot{f}_\theta^{(j)}(u_j|u_{j-1}) \prod_{t=j+1}^{T} f_\theta^{(t)}(u_t|u_{t-1})\, d\nu(u_1) \ldots d\nu(u_T)$$

under the assumptions

Assumptions.

- (i) The transition densities satisfy

$$\sup_{\theta\in\Theta} \|f_\theta^{(t)}(\cdot|\cdot)\|_{p,\infty} < \infty \quad \forall t$$

- (ii) The transition densities are differentiable in the following sense: There are functions $\dot{f}_\theta^{(t)}(\cdot|\cdot)$ such that
$$||\dot{f}_\theta^{(t)}(.|.)||_{p,\infty} < \infty$$
$$\|\frac{1}{h}[f_{\theta+h}^{(t)}(\cdot|\cdot) - f_\theta^{(t)}(\cdot|\cdot)] - \dot{f}_\theta^{(t)}(\cdot|\cdot)\|_{p,\infty} \to 0$$
for $h \to 0$.

- (iii) There is a function $v \in L_q(d\nu)$ $(1/p + 1/q = 1)$, such that
$$|\overline{H}(u_1, \ldots, u_T)| \leq C \prod_{t=1}^{T} v(u_t).$$

The proof of 4.6.17 follows the principle of the proof of 4.6.15.

Example (Lookback-Option in the Black Scholes model: ρ (homogeneous case))**.** The payoff function for the Lookback-Option is of the following form
$$H(S(T)) = e^{-rT} \max\left\{\max_{0\leq t\leq T}(S(t)) - K, 0\right\}$$
which is written as a function of the driving process as (see equation 4.6.12)
$$\overline{H}(u_1, ..., u_T) = e^{-rT} \max\left\{\max_{0\leq t\leq T}(e^{u_1}, ..., e^{u_T}) - K, 0\right\}$$
with $S(0) = 1$ and $u_{t+1} = u_t + N(\mu, \sigma^2) = N(\mu + u_t, \sigma^2)$. The parameter of interest is μ.
Therefore the transition densities are
$$f_\mu(u_t|u_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(u_t - u_{t-1} - \mu)^2}{2\sigma^2}\right).$$

Our aim is to find a measure $\nu(u)$, more precisely we are seeking for a density $\psi(u)$ with $d\nu(u) = \psi(u)du$, and a function $v(u_t)$, such that assumption i),ii),iii) and therefore 4.6.14 holds. Formula 4.6.14 is then of the form
$$|\int \cdots \int \prod_{t=1}^{T} v(u_t) \cdot \frac{f_\mu(u_t|u_{t-1})}{\psi(u_t)} \psi(u_1)du_1 \ldots \psi(u_T)du_T| \leq (\|v\|_q)^T \cdot \prod_{t=1}^{T} \|\frac{f_\mu(\cdot|\cdot)}{\psi(\cdot)}\|_{p,\infty}. \tag{4.6.18}$$

First let us turn toward assumption iii.). An upper bound for the payoff can be

$$\begin{aligned}\overline{H}(u_1,...,u_T) &= e^{-rT}\max\left[\max\left(e^{u_1},...,e^{u_T}\right)-K,0\right] \\ &\le e^{-rT}e^{\max(u_1,...,u_T)} \\ &\le e^{-rT}\prod_{t=1}^{T}\underbrace{e^{u_t}}_{v(u_t)}\end{aligned}$$

and for the density $\psi(u_t)$ we take

$$\psi(u_t)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{u_t^2}{2\sigma^2}\right).$$

So we have to show that $e^{u_t}\in L^q(\psi(u_t)du_t)$. This is the case because of

$$\int(e^{u_t})^q\psi(u_t)du_t=\frac{1}{\sqrt{2\pi}\sigma}\int e^{u_tq}\cdot e^{-\frac{u_t^2}{2\sigma^2}}du_t<\infty.$$

Now let

$$\varphi(u_t|u_{t-1}):=\frac{f_\mu(u_t|u_{t-1})}{\psi(u_t)}.$$

For assumption ii.) we show that $\frac{\partial}{\partial\mu}\frac{f_\mu(u_t|u_{t-1})}{\psi(u_t)}=\dot{\varphi}(u_t|u_{t-1})$ exists and that it is in $L^p(\psi(u_t)du_t)$, and by Theorem 7 (see Appendix) we get the rest. So we have

$$\varphi(u_t|u_{t-1})=\exp\left(\frac{u_t(u_{t-1}+\mu)}{\sigma^2}-\frac{(u_{t-1}+\mu)^2}{2\sigma^2}\right)\in L^p(\psi(u)du.$$

Let $\sigma=1$ then we get for the derivative

$$\dot{\varphi}(u_t|u_{t-1})=\exp\left(u_t(u_{t-1}+\mu)-\frac{(u_{t-1}+\mu)^2}{2}\right)\cdot(u_t-u_{t-1}-\mu).$$

Let $(u_{t-1}+\mu)=z$ so we have to proof that $\exp\left(u_tz-\frac{z^2}{2}\right)\cdot(u_t-z)\in L^p(\psi(u_t)du_t)$. Therefore we have

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}}\int\exp\left(u_tz-\frac{z^2}{2}\right)^p\cdot(u_t-z)^p\exp\left(-\frac{u_t^2}{2}\right)du_t \\ =\;&\frac{1}{\sqrt{2\pi}}\int\exp\left(p\cdot z\cdot u_t-\frac{z^2p}{2}-\frac{u_t^2}{2}\right)(u_t-z)^pdu_t \\ =\;&\frac{1}{\sqrt{2\pi}}\int\exp\left(-\frac{(u_t-p\cdot z)^2}{2}+\frac{p^2z^2}{2}-\frac{z^2p}{2}\right)(u_t-z)^pdu_t \\ =\;&\exp\left(\frac{p^2z^2}{2}-\frac{z^2p}{2}\right)\underbrace{\frac{1}{\sqrt{2\pi}}\int\exp\left(-\frac{(u_t-p\cdot z)^2}{2}\right)(u_t-z)^pdu_t}_{=1}<\infty\quad\forall p<\infty.\end{aligned}$$

By Theorem 7 (see Appendix) it follows that $\dot{\varphi}(.|.)$ is the L^p-derivative of $\varphi(.|.)$ and assumption ii.) is satisfied.
So for Lookback Options in the Black-Scholes case, we found a densityquotient $\psi(u_t)$, where assumptions i.), ii.) and iii.) are satisfied. For options different to Lookback Options, the only challenge we face is to find an upper bound for the payoff function, such that assumption iii.) is satisfied. This can be usually done very quickly.

Example (Asian Option in the Black Scholes model: ρ (homogeneous case))**.** The Asian Option is of the following form

$$H(S(T)) = e^{-rT} \max\left[\frac{1}{T}\sum_{i=1}^{T} S(i) - K, 0\right]$$

or

$$\overline{H}(u_1, ..., u_T) = e^{-rT} \max\left[\frac{1}{T}\sum_{i=1}^{T} e^{u_i} - K, 0\right]$$

with $u_{t+1} = u_t + N(\mu, \sigma^2) = N(\mu + u_t, \sigma^2)$. The parameter of interest is μ.

We just have to concentrate on assumption iii.), because the proof of assumption i.) and ii.) follows the principle of the previous example. An upper bound for the payoff can be

$$\begin{aligned}
\overline{H}(u_1, ..., u_T) &= e^{-rT} \max\left[\frac{1}{T}\sum_{i=1}^{T} e^{u_i} - K, 0\right] \\
&\leq e^{-rT}\frac{1}{T}\sum_{i=1}^{T} e^{u_i} \\
&\leq e^{-rT}\frac{1}{T}\prod_{t=1}^{T}\underbrace{(1 + e^{u_t})}_{v(u_t)}
\end{aligned}$$

Recall that

$$\psi(u_t) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{u_t^2}{2\sigma^2}\right).$$

So we have to show that $(1 + e^{u_t}) \in L^q(\psi(u_t)du_t)$. This is the case because of

$$\int (1 + e^{u_t})^q \psi(u_t) du_t \leq \frac{1}{\sqrt{2\pi}\sigma}\int 2^q(1 + e^{u_t q}) \cdot e^{-\frac{u_t^2}{2\sigma^2}} du_t < \infty.$$

4.6.8.2 Estimation via Operators

We can regard the problem also from a different point of view, by constructing operators corresponding to the maximum respectively average of the stock.

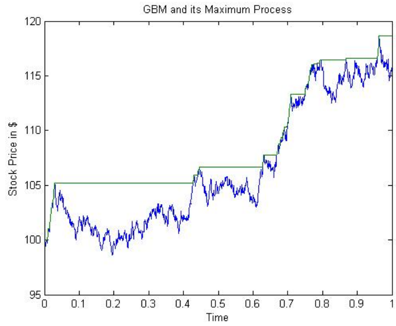


Figure 4.6.4: Example: Maximums-process.

Lookback Options A Lookback Option is based on the maximum value of the underlying. Let $S_\theta(\cdot)$ be the Markovian price process of the underlying. Define the maximum process as $M_\theta(i) = \max_{k\leq i} S_\theta(k)$. Recall that the maturity time T corresponds to the step n with $T = n \cdot \Delta$. The payoff of a Lookback Option is $H(M_\theta(n))$ and its fair price is $\mathbb{E}[e^{-rT} H(M_\theta(n))]$. Notice that the maximum process M_θ is not Markovian itself, but it is the first component of the Markovian pair $(M_\theta(\cdot), S_\theta(\cdot))$. While the evolution of the second component is as before and given by the transition of $X_\theta = \log(S(t))$, the transition of the first components is given by

$$\begin{aligned} M_\theta(1) &= S_\theta(1) \\ M_\theta(i+1) &= \begin{cases} S_\theta(i+1) & \text{if } S_\theta(i+1) > M_\theta(i) \\ M_\theta(i) & \text{otherwise} \end{cases} \end{aligned}$$

Let $\mathbb{P}_\theta^{(M)}$ be the transition operator of this two-dimensional process. We aim an unbiasedly estimate

$$\frac{\partial}{\partial\theta}\mathbb{E}[e^{-rT} H(M_\theta(n))] = \frac{\partial}{\partial\theta} e^{-rT}\gamma \left(\mathbb{P}_\theta^{(M)}\right)^n H$$

Here - of course - H applies only to the first component of the Markov process.

The MVDr algorithm for the Lookback Option.

1. Sample random uniform time τ in $\{1, ..., n\}$. Sample $S_\theta(1), ..., S_\theta(\tau)$ and calculate $M_\theta(1), \ldots, M_\theta(\tau)$.

2. Then at time τ do
Use the positive resp. negative part of the derivative of $\mathbb{P}_\theta$ to make one transition from $S_\theta(\tau - 1)$ to $S_\theta^+(\tau)$ and $S_\theta^-(\tau)$ respectively. Store c_θ.
Let $M_\theta^+(\tau) = \max(M(\tau-1), S_\theta^+(\tau))$ and $M_\theta^-(\tau) = \max(M(\tau-1), S_\theta^-(\tau))$.

3. Continue the processes $S_\theta^+(\cdot)$ and $S_\theta^-(\cdot)$ using the transition $\mathbb{P}_\theta$ and calculate $(M_\theta^+(\ell), S_\theta^+(\ell))$ resp. $(M_\theta^-(\ell), S_\theta^-(\ell)$ for $\ell = \tau + 1, \ldots, n$ with

$$M_\theta^+(\ell) = \begin{cases} S_\theta^+(\ell) & \text{if } S_\theta^+(\ell) > M_\theta^+(\ell-1) \\ M_\theta^+(\ell-1) & \text{otherwise} \end{cases}$$

resp.

$$M_\theta^-(\ell) = \begin{cases} S_\theta^-(\ell) & \text{if } S_\theta^-(\ell) > M_\theta^-(\ell-1) \\ M_\theta^-(\ell-1) & \text{otherwise} \end{cases}$$

4. The unbiased estimate is

$$n \cdot c_\theta \cdot [H(M_\theta^+(n)) - H(M_\theta^-(n))].$$

5. The final estimate is the arithmetic mean of N independent replications of estimate 4.

To see the correctness of the algorithm notice that $\mathbb{T}(x) = \max(c, x)$ is continuous therefore by Remark 1 it may be composed with the weak derivative pair.

As before, the MVDr algorithm constructs just one phantom pair. In complete analogy to the standard Option case, the MVDe resp. the MVDc(k) estimates can be constructed based on n resp. k pairs of phantoms.

Asian Options An Asian Option is based on the average value of the underlying. Let $S_\theta(\cdot)$ be the Markovian price process of the underlying. Define the average process as $A_\theta(k) = \frac{1}{k}\sum_{i=1}^k S_\theta(i)$. The payoff of a Lookback Option is $H(A_\theta(n))$ and its fair price is $\mathbb{E}[e^{-rT}H(A_\theta(n))]$. Notice that the average process A_θ is not Markovian itself, but it is the first component of the Markovian pair $(A_\theta(\cdot), S_\theta(\cdot))$. While the evolution of the second component is as before and given by the transition of $X_\theta = \log(S(t))$, the transition of the first components is given by

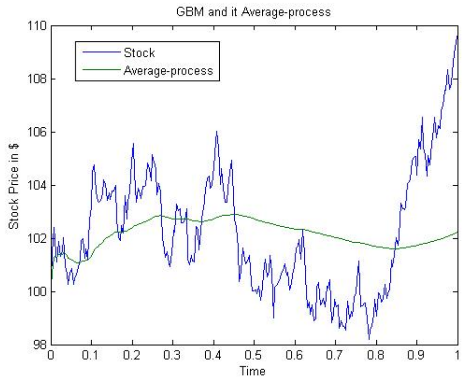
$$A_\theta(i) = \frac{i}{i+1}A_\theta(i-1) + \frac{1}{i+1}S_\theta(i) \tag{4.6.19}$$

Let $\mathbb{P}_\theta^{(A)}$ be the transition operator of this two-dimensional process. We aim an unbiasedly estimate

$$\frac{\partial}{\partial\theta}\mathbb{E}[e^{-rT}H(A_\theta(n))] = \frac{\partial}{\partial\theta}e^{-rT}\gamma\left(\mathbb{P}_\theta^{(A)}\right)^n H$$

. Here - of course - H applies only to the first component of the Markov process.

The following figure shows an example for the average process, when $T = 1$ (year) with 252 trading days.



The MVDr algorithm for the Asian Option.

1. Sample random uniform time τ in $\{1, ..., n\}$. Sample $S_\theta(1), ..., S_\theta(\tau-1)$ and calculate $A_\theta(1), \ldots, A_\theta(\tau-1)$.

2. Then at time τ do: Use the positive resp. negative part of the derivative of $\mathbb{P}_\theta$ to make one transition from $S_\theta(\tau-1)$ to $S^+_\theta(\tau)$ and $S^-_\theta(\tau)$ respectively. Store c_θ. Let

$$A^+_\theta(\tau) = \frac{1}{\tau+1}\left(\underbrace{\tau \cdot A(\tau-1)}_{\sum_{i=0}^{\tau-1} S_\theta(i)} + S^+_\theta(\tau)\right)$$

resp.

$$A^-_\theta(\tau) = \frac{1}{\tau+1}\left(\underbrace{\tau \cdot A_\theta(\tau-1)}_{\sum_{i=0}^{\tau-1} S_\theta(i)} + S^-_\theta(\tau)\right).$$

3. Continue the processes $S^+_\theta(\cdot)$ and $S^-_\theta(\cdot)$ using the transition $\mathbb{P}_\theta$ and calculate $(A^+_\theta(\ell), S^+_\theta(\ell))$ resp. $(A^-_\theta(\ell), S^-_\theta(\ell)$ for $\ell = \tau+1, \ldots, n$ with

$$A^+_\theta(l) = \frac{l}{l+1} \cdot A^+_\theta(l-1) + \frac{1}{l+1} S^+_\theta(l)$$

resp.

$$A^-_\theta(l) = \frac{l}{l+1} \cdot A^-_\theta(l) + \frac{1}{l+1} S^-_\theta(l).$$

4. The unbiased estimate is

$$n \cdot c_\theta \left[H(A^+_\theta(n)) - H(A^-_\theta(n))\right]. \tag{4.6.20}$$

5. The final estimate is the arithmetic mean of N independent replications of estimate (4.6.20).

Also in this case, one may use the variants MVDe and MVDc(k).

5 Implementation of MVD for exponential Lévy processes

In this section we present numerical results of sensitivity estimates where we consider discrete time homogeneous Markov processes with infinitely divisible increment distribution. Recall that the fair option price is $\mathbb{E}[e^{-rT}H(S_\theta(n))]$ and we are interested in estimating

$$\frac{\partial}{\partial\theta}\mathbb{E}[e^{-rT}H(S_\theta(n))] = \frac{\partial}{\partial\theta}e^{-rT}\gamma\,\mathbb{P}_\theta^n\,H \tag{5.0.1}$$

Remark 6. Notice that in case that the derivative is taken w.r.t. the interest rate r, i.e. if $\theta = r$, also the derivative w.r.t. the discount factor e^{-rT} has to be taken into account. In this case

$$\frac{\partial}{\partial r}\mathbb{E}[e^{-rT}H(S_r(n))] = -Te^{-rT}\mathbb{E}[H(S_r(n))] + e^{-rT}\frac{\partial}{\partial r}\gamma\,\mathbb{P}_r^n\,H. \tag{5.0.2}$$

If however the sensitivity w.r.t. some other parameter is searched for, we have just to calculate

$$\frac{\partial}{\partial\theta}\mathbb{E}[e^{-rT}H(S_\theta(n))] = e^{-rT}\frac{\partial}{\partial\theta}\gamma\,\mathbb{P}_\theta^n\,H. \tag{5.0.3}$$

In all the numerical examples we show, the time unit is 1 year and the maturity time was set to $T = 1$. The elementary time step was set to one trading day thus we chose $n = 252$ and $\Delta = 1/n$.

5.1 Geometric Brownian motion model

5.1.1 Rho (ρ): Sensitivity w. r. t. r

$$\rho = \frac{\partial}{\partial r}\mathbb{E}\left[e^{-rT}H(S_r(T))\right].$$

Renaming the process $S_\theta(i\Delta)$ by $S_\theta(i)$ we arrive at

$$\log S(i+1) - \log S(i) \sim N\left(\left(r - \frac{1}{2}\sigma^2\right)\frac{1}{n}, \sigma^2\frac{1}{n}\right)$$

According to the weak derivative triplet of a Normal distribution, (see Table 4.1), the positive resp. negative realizations of the weak derivative are

$$\begin{aligned}
[\log S(i+1) - \log S(i)]^+ &= (r - \sigma^2/2)/n + V^+ \\
[\log S(i+1) - \log S(i)]^- &= (r - \sigma^2/2)/n - V^-
\end{aligned}$$

with

$$V^{+} = V^{-} \sim \text{Rayleigh}(\sigma^2/n)$$

and

$$c_r = \frac{1}{\sigma\sqrt{2\pi n}}.$$

Example (The Rho (ρ) for a Plain Vanilla Option)**.** We consider now the payoff function

$$H(S(T)) = e^{-rT}\max\{S(T) - K, 0\}.$$

It is well known that the exact value Rho of an option on a non-dividend-paying stock is defined as

$$\rho = KTe^{-rT}\Phi(d_2)$$

with

$$d_2 = \frac{ln(S(0)/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}$$

where $\Phi(x)$ is the standard normal cumulative probability function, see [7, p.337].

Table 5.1 shows the performance of IPA (=Pathwise method), the FD (=Finite Difference method), MVDr (=randomized phantom estimator), MVDe (=Measure Valued Differentitiation with total differentiation) and MVDc(k) (Measure Valued Differentiation with differentiation at k random points) when calculating the ρ for a plain Vanilla Option, whereas the Measure Valued Derivatives were computed according to 5.0.2.
One simulation of each estimator is based on $N = 200$ replications of the path of the stock. [1] For FD $\hat{\Delta}_{FD} = \frac{H(S_{r+l}(T) - H(S_r(T))}{l}$, $l = 0.01$. Further $S(0) = 100$, $r = 0.01, \sigma = 0.05, K = 100$. To compare the computational effort of the methods we use the "work-normalized variance"(WNV) which is given by the product of the variance and the expected work per run, [14].

[1] Throughout the paper each estimator was simulated 300 times to get the arithmetic mean of the estimator, the arithmetic mean of the computational time and the Variance

	$\widehat{\rho}$	Variance	computational time	WNV
exact value	7.15			
MVDr	7,070	4,508	0,783	3,528
MVDe	7,135	2,015	4,459	8,986
MVDc(6)	7,162	2,752	0,869	2,391
FD	6,37	23,773	1,5298	36,368

Table 5.2: The Rho for a Digital Option in the BS-model, $N = 200$

	$\widehat{\rho}$	Variance	computational time	WNV
exact value	56,379			
IPA	56,058	14,120	0,778	10,992
MVDr	55,961	18,159	0,793	14,407
MVDe	56,673	11,051	3,533	39,036
MVDc(2)	56,180	16,684	0,795	13,177
MVDc(4)	56,243	14,589	0,855	12,469
MVDc(6)	56,393	13,359	0,923	12,330
MVDc(8)	56,625	13,059	0,969	12,649
MVDc(16)	56,258	11,089	1,201	13,318
FD	63,135	1006,390	1,574	1584,51

Table 5.1: The Rho for a Vanilla Option in the BS-model, $N = 200$.

When comparing the results we see that the IPA method performs better than the MVD algorithms. If we compare the MVD algorithms among each other we observe, that the MVDc has the smallest work normalized variance. As it could be expected it lies in between the MVD and MVDe in terms of variance and computational time.

Example (The Rho (ρ) for a Digital Option)**.**

The Payoff function of a Digital Option is

$$H(S(T)) = e^{-rT}\mathbf{1}\{S(T) > K\}.$$

With the same inputs as we had before for the Vanilla Option, we obtain results shown in Table 5.2 for the ρ of the Digital Option. The value of BS-Rho was taken out of www.mathfinance.de/optioncalculator.php. Notice that it IPA cannot be applied in the case of Digital Options due to the fact the the payoff is not continuous at the point of non-differentiability.

Example (The Rho (ρ) for a Lookback Option)**.**

The payoff function of an Lookback Option is defined as

$$H(S(T)) = e^{-rT}[\sup_{0\leq t\leq T} S(t) - K]^{+}$$

Estimation via joint densities: The randomized phantom estimator (MVDr) is

$$\begin{aligned}\frac{\partial \mathbb{E}\left(H(S(T))\right)}{\partial r} &= \mathbb{E}\left[-Te^{-rT}\max\left\{\sup_{0\leq t\leq T} S(t) - K, 0\right\}\right.\\ &+ \left. Te^{-rT}\cdot c\cdot n\cdot\left(\max\left\{\sup_{0\leq t\leq T} S(t)^{+} - K, 0\right\} - \max\left\{\sup_{0\leq t\leq T} S(t)^{-} - K, 0\right\}\right)\right]\end{aligned} \tag{5.1.1}$$

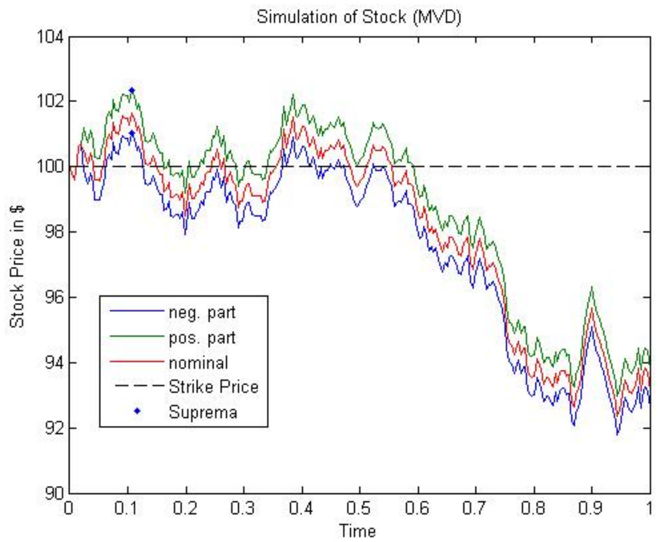


Figure 5.1.1: Lookback Option: Example of generated phantoms.

Let $S(0) = 100$, $r = 0.01, \sigma = 0.05, K = 100, T = 1$ and the number of trading days is 252. Within this framework we get results shown in Table 5.3, whereas the value for Rho was taken from www.wolframalpha.com.

	Estimator	Variance	computational time	WNV
BS-Rho	52,90			
MVDr	52,849	19,557	0,761	14,873
MVDe	52,933	5,236	5,620	29,430
MVDc(6)	52,947	6,481	0,853	5,528

Table 5.4: The Rho for a Lookback Option in the BS-model via the Maximums Process , $N = 200$.

	Estimator	meanVariance	computational time	WNV
Rho	52,90			
MVD	52,923	20,277	0,788	15,987
MVDe	52,861	5,421	5,630	30,520
MVDc	52,963	6,606	0,849	5,608

Table 5.3: The Rho for a Lookback Option in the BS-model , $N = 200$.

Estimation via Maximums Operator

$$\begin{aligned}\frac{\partial \mathbb{E}[H(S(T))]}{\partial r} &= \mathbb{E}\left[-T \cdot e^{-r \cdot T} \cdot \max\{M(T) - K, 0\}\right. \\ &+ \left. e^{-r \cdot T} \cdot n \cdot T \cdot c \cdot \left(\max\{M^+(T) - K, 0\} - \max\{M^-(T) - K, 0\}\right)\right].\end{aligned}$$

Consider the same input values as before, we get results which are stated in Table 5.4. Notice that the results are almost the same as in Table 5.3. That leads to the conclusion that for Lookback-Options, the Esimator via the Maximums-process is as efficient as the Estimator via equation (5.1.1). The disadvantage stems from the fact, that with this kind of compuatation, we have to find the transition operator for every different payoff function. Whereas if Lemma 4.6.8.1 can be applied[2], we just have to find an upper polynomial bound for the payoff function.

Example (The Rho (ρ) for an Asian Option)**.**

The payoff function of an Asian Call Option is defined as

$$H(S_T) = e^{-rT} \max\left\{0, \frac{1}{T}\sum_{i=1}^{T} S(T) - K\right\}.$$

[2]if an appropriate density quotient can be found

	Estimator	meanVariance	computational time	WNV
BS-Rho	27.111			
MVDr	27,38	7,95	0,7861	6,25
MVDe	27,12	2,49	5,4087	13,67
MVDc	27,44	2,82	0,8476	2,39

Table 5.5: The Rho for an Asian Option in the BS-model , $N = 200$.

Estimation via joint distribution: The randomized phantom estimator (MVDr) is

$$\begin{aligned}\frac{\partial \mathbb{E}\left(H(S_T)\right)}{\partial r} &= \mathbb{E}\left[-Te^{-rT}\max\left\{0, \frac{1}{T}\sum_{i=1}^{T} S(T) - K\right\}\right. \\ &+ \left. Te^{-rT}\cdot c\cdot n\cdot\left(\max\left\{0, \frac{1}{T}\sum_{i=1}^{T} S(T)^{+} - K\right\} - \max\left\{0, \frac{1}{T}\sum_{i=1}^{T} S(T)^{-} - K\right\}\right)\right]\end{aligned} \tag{5.1.2}$$

In an simulation, with $S(0) = 100$, $r = 0.01, \sigma = 0.05, K = 100, T = 1$ we get results shown in Table 5.5, whereas the value for BS-Rho was taken from www.wolframalpha.com.

Estimation via Average Operator: The final estimate is

$$\begin{aligned}\frac{\partial \mathbb{E}[H(S(T))]}{\partial r} &= \mathbb{E}\left[-T\cdot e^{-r\cdot T}\cdot\max\{0, A(T) - K\}\right. \\ &+ \left. e^{-r\cdot T}\cdot T\cdot n\cdot c\cdot\left(\max\left\{0, A(T)^{+} - K\right\} - \max\left\{0, A(T)^{-} - K\right\}\right)\right].\end{aligned}$$

With $S(0) = 100$, $r = 0.01, \sigma = 0.05, K = 100, T = 1$ we get results shown in Table 5.6. Comparing with Table 5.5, we can see that also here the estimator via the Average-process leads to nearly the same results as the estimator via equation (5.1.2), and therefore is as efficient.

5.1.2 Vega (ν): Sensitivities w.r.t. σ

$$\nu = \frac{\partial}{\partial\sigma}\mathbb{E}\left[r^{-rT}H(S_\sigma(T))\right]$$

The log returns follow

$$\log S(i+1) - \log S(i) = V$$

	Estimator	Variance	computational time	WNV
BS-Rho	27.111			
MVDr	27,30	8,12	0,7617	6,18
MVDe	27,14	2,80	6,3196	17,68
MVDc(6)	26,99	3,01	0,8968	2,70

Table 5.6: The Rho for an Asian Option in the BS-model via the Average process, $N = 200$.

with

$$V \sim N\left(\underbrace{\left(r - \frac{1}{2}\sigma^2\right)\frac{1}{n}}_{\mu_\sigma}, \underbrace{\sigma^2\frac{1}{n}}_{\sigma_\sigma{}^2}\right).$$

According to Example 4.6.2 on page 50, the positiv and negative part are with probability p

$$\begin{aligned}
[\log S(i+1) - \log S(i)]^+ =& (r - \sigma^2/2)/n - V^+ \\
[\log S(i+1) - \log S(i)]^- =& (r - \sigma^2/2)/n + V^-
\end{aligned}$$

and with probability $1 - p$

$$\begin{aligned}
[\log S(t_{i+1}) - \log S(t_i)]^+ &= (r - \sigma^2/2)/n + W \\
[\log S(t_{i+1}) - \log S(t_i)]^- &= (r - \sigma^2/2)/n + U \cdot W
\end{aligned}$$

with

$$\begin{aligned}
V^+ &= V^- \sim \text{Rayleigh}(\sigma^2/n) \\
W &\sim \text{ds-Maxwell}(0, \sigma^2/n) \\
U &\sim \text{Uniform}(0,1)
\end{aligned}$$

whereas

$$p = \frac{\frac{|\mu'_\sigma|}{\sigma_\sigma\sqrt{2\pi}}}{\frac{|\mu'_\sigma|}{\sigma_\sigma\sqrt{2\pi}} + \frac{\sigma'_\sigma}{\sigma_\sigma}} = \frac{|\mu'_\sigma|}{|\mu'_\sigma| + \sigma'_\sigma\sqrt{2\pi}} = \frac{\frac{\sigma}{n}}{\frac{\sigma}{n} + \sqrt{\frac{2\pi}{n}}}.$$

For c_σ we get according to 4.6.2

$$c_\sigma = |\mu'_\sigma| \frac{1}{\sigma_\sigma\sqrt{2\pi}} + \frac{\sigma'_\sigma}{\sigma_\sigma} = \frac{|\mu'_\sigma| + \sigma'_\sigma\sqrt{2\pi}}{\sigma_\sigma\sqrt{2\pi}} = \frac{\frac{\sigma}{n} + \sqrt{\frac{2\pi}{n}}}{\sigma\sqrt{\frac{2\pi}{n}}}.$$

Example (The ν for a Plain Vanilla Option:).

	$\widehat{\nu}$	meanVariance	computational time	WNV
exact value	38,897			
IPA	38,923	19,393	0,789	15,293
MVDr	39,031	6889,428	0,771	5311,487
MVDe	38,715	83,938	5,635	473,001
MVDc(6)	38,943	1328,352	0,862	1144,891
FD	38,428	1186,400	1,5324	1818,002

Table 5.7: The Vega for a Vanilla Option in the BS-model, $N = 200$.

The exact value for Vega for an European Call or Put Option on a non-dividend-paying stock is defined by

$$\nu = S(0)\sqrt{T}\Phi(d_1)$$

with

$$d_1 = \frac{\ln(S(0)/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

see [7, p.336].

With the input arguments of the previous examples, we get results shown in Table 5.7.

5.1.3 Delta (Δ): Sensitivities w.r.t. $S(0)$

$$\Delta = \frac{\partial}{\partial S(0)} \mathbb{E}\left[r^{-rT} H(S_{S(0)}(T))\right]$$

$$\log S(1) - \log S(0) = V$$

with

$$V \sim N\left(\left(r - \frac{1}{2}\sigma^2\right)\frac{1}{n}, \sigma^2 \frac{1}{n}\right).$$

According to the weak derivative of a Normal distribution, see 4.1, the process splits up into

$$\begin{aligned} [\log S(1) - \log S(0)]^+ &= (r - \sigma^2/2)/n + V^+ \\ [\log S(1) - \log S(0)]^- &= (r - \sigma^2/2)/n - V^- \end{aligned}$$

with

$$V^+ = V^- \sim \text{Rayleigh}(\sigma^2/n)$$

	$\widehat{\Delta}$	meanVariance	computational time	WNV
exact value	0,589	0,000	0,000	0,000
IPA	0,588	0,0013	0,7819	0,0010
MVDr	0,594	0,0022	0,7612	0,0017
FD	0,950	1122,625	1,498	1682,507

Table 5.8: The Delta for a Vanilla Option in the BS-model, $N = 200$.

and

$$c_{S(0)} = \frac{1}{S(0)\sigma\sqrt{2\pi\frac{1}{n}}}.$$

Example (The Δ for a Plain Vanilla Option:)**.**

Consider the payoff function

$$H(S_T) = e^{-rT} \max\left\{S(T) - K, 0\right\}.$$

The exact value of Delta of an option on a non-dividend-paying stock is defined as

$$\Delta = \Phi(d_1)$$

with

$$d_1 = \frac{ln(S(0)/K) + (r + \sigma/2)T}{\sigma\sqrt{T}}$$

where $\Phi(x)$ is the cumulative probability function for a standard normal variable, see [7].
With the same inputs as we had before while calculating the ρ, we obtain results shown in Table 5.8 for the Delta (Δ) of the Vanilla Option.

Example (The Delta (Δ) for Digital Options)**.**

With the same inputs as we had before for the Vanilla Option, we obtain results shown in Table 5.9 for the Δ of the Digital Option. The value of BS-Delta was taken out of www.mathfinance.de/optioncalculator.php.

	$\widehat{\Delta}$	meanVariance	computational time	WNV
Delta	$0,078$			
MVDr	0,078	0,0005	0,760	0,0004
FD	-0,238	22,844	1,519	34,707

Table 5.9: The Delta for a Digital Option in the BS-model, $N = 200$.

5.2 Poisson model

Sensitivity w.r.t. λ

The log returns are distributed like

$$\log S(i+1) - \log S(i) = V$$

with

$$V \sim P\left(\lambda \frac{1}{n}\right).$$

The positive resp. negative phantoms are

$$\begin{aligned}
[\log S(i+1) - \log S(i)]^+ &= V + 1 \\
[\log S(i+1) - \log S(i)]^- &= V \\
c_\lambda &= \frac{1}{n}
\end{aligned}$$

To get appropriate values, we calibrate by $e^{1/d}$ when simulating the stock, so

$$S(i+1) = S(i) e^{1/d \cdot Poisson(\frac{1}{n}\lambda)}.$$

Example (λ for Plain Vanilla Option)**.**

$$H(S_T) = e^{-rT} \max\{S(T) - K, 0\}.$$

MVDr:

$$\frac{\partial \mathbb{E}\left(H(S_T)\right)}{\partial r} = \mathbb{E}\left[e^{-rT} \cdot c_\lambda \cdot T \cdot n \cdot \left(\max\{S(T)^+ - K, 0\} - \max\{S(T)^- - K, 0\}\right)\right]$$

With input parameters $S(0) = 100$, $r = 0.01$, $\lambda = 10, l = 0.01$, $K = 110$, $d = 100$, and $h = 1$ we get results shown in Table 5.10.

Example (λ for a Digital Option)**.**

$$H(S_T) = e^{-rT} \mathbf{1}\{S(T) > K\}$$

With the same input options as above we get results shown in Table 5.11.

	Estimator	Variance	computational time	WNV
MVDr	0,675	0,001689	6,50	0,01098
MVDe	0,677	0,001415	7,54	0,01067
MVDc(6)	0,679	0,001694	6,94	0,01176
FD	3,764	634,83	12,42	7884,59

Table 5.10: Sensitivity w.r.t. λ of a Vanilla Option in the Poisson model, $N = 200$

	Estimator	Variance	computational time	WNV
MVDr	0,6573	0,0012	6,14	0,0074
MVDe	0,6581	0,0010	6,84	0,0072
MVDc(6)	0,6571	0,0011	6,25	0,0068
FD	0,3878	24,86	12,73	316,47

Table 5.11: Sensitivity w.r.t. λ of a Digital Option in the Poisson model, $N = 200$

5.3 Gamma model

Sensitivity w.r.t. b

$$\log S(i+1) - \log S(i) = V$$

with

$$V \sim \text{Gamma}\left(\frac{1}{n}a, b\right)$$

Notice that

$$\begin{aligned}
\frac{\partial}{\partial\theta}(1-iub)^{-\frac{a}{n}} &= iu\frac{a}{n}(1-iub)^{-\frac{a}{n}-1} \\
&= \frac{a}{nb}\left(iub(1-iub)^{-(\frac{a}{n}+1)}\right) \\
&= \frac{a}{nb}\left((1-iub)^{-(\frac{a}{n}+1)} - (1-iub)(1-iub)^{-(\frac{a}{n}+1)}\right) \\
&= \frac{a}{nb}\left(\underbrace{(1-iub)^{-(\frac{a}{n}+1)}}_{\sim Gamma(\frac{a}{n}+1,b)} - \underbrace{(1-iub)^{-\frac{a}{n}}}_{\sim Gamma(\frac{a}{n},b)}\right)
\end{aligned}$$

therefore

$$\begin{aligned}
[\log S(i+1) - \log S(i)]^{+} &= V^{+} \\
[\log S(i+1) - \log S(i)]^{-} &= V
\end{aligned}$$

with

$$V^{+} \sim \text{Gamma}\left(\frac{1}{n}a+1, b\right)$$

	Estimator	Variance	computational time	WNV
MVDr	1,401	0,019	5,12	0,096
MVDe	1,398	0,002	11,19	0,026
FD	1,43	125,23	9,45	1183,42

Table 5.12: Sensitivity w.r.t. b of a Vanilla Option in the Gamma model, $N = 200$

and

$$c_b = \frac{a}{bn}.$$

To get appropriate values, we calibrate by $e^{1/d}$ when simulating the stock, so

$$S(i+1) = S(i)e^{(1/d)\cdot \text{Gamma}(\frac{1}{n}a,b)}.$$

Example (b for Plain Vanilla Option)**.**

Input parameters: $S(0) = 100$, $r = 0.01, a = 2, b = 1$, $l = 0.01$, $K = 102$, $h = 1, d = 100$.

5.4 The Variance Gamma model

Sensitivity w.r.t. b_1

$$\frac{\partial}{\partial b_1}\mathbb{E}\left[r^{-rT}H(S_{b_1}(T))\right]$$

The log returns follow

$$\log S(i+1) - \log S(i) = V_1 - V_2$$

with

$$V_1 \sim \text{Gamma}\left(\frac{a}{n}, b_1\right)$$
$$V_2 \sim \text{Gamma}\left(\frac{a}{n}, b_2\right)$$

and

$$\begin{aligned}[\log S(i+1) - \log S(i)]^+ &= V_1^+ - V_2 \\ [\log S(i+1) - \log S(i)]^- &= V_1 - V_2\end{aligned}$$

with

$$V_1^+ \sim \text{Gamma}(\frac{a}{n} + 1, b_1).$$

Further

$$c_{b_1} = \frac{a}{b_1 n}.$$

	Estimator	Variance	computational time	WNV
MVDr	76,08	45,30	9,79	443,63
MVDe	75,42	5,08	20,55	104,36
FD	85,99	253,97	18,44	4683,21

Table 5.13: Sensitivity w.r.t. b_1 of a Vanilla Option in the VG-model, $N = 200$.

Example (Sensitivity w.r.t. b_1 for a Plain Vanilla Call Option).

In a simulation with input parameters: $r = 0.01, a = 1, b_1 = 0.01, b_2 = 0.01, d = 100, \text{start} = 100, K = 100, n = 252, l = 0,01$ we get results shown in Table 5.13 .

Example (Sensitivity w.r.t. b_1 for a Lookback Option).

Recall the payoff function

$$H(S(T)) = e^{-rT}[\sup_{0\leq t\leq T} S(t) - K]^+.$$

In the case of path dependent payoff functions in the Variance Gamma model, Lemma 4.6.8.1 cannot be applied in a straightforward way. The problem is to find the appropriate density quotient $\psi(u_t)$ in

$$|\int \dots \int \prod_{t=1}^{T} v(u_t) \cdot \frac{f(u_t|u_{t-1})}{\psi(u_t)} \psi(u_1)du_1 \dots \psi(u_T)du_T| \leq (\|v\|_q)^T \cdot \prod_{t=1}^{T} \|\frac{f(\cdot|\cdot)}{\psi(\cdot)}\|_{p,\infty}$$

because in the conditional density $f(u_t|u_{t-1})$ of the Variance Gamma process, there occur Bessel functions of the second kind, see [2].
At that point, we can take advantage of considering the Maxiums-process as a Markov process, as we also did for Lookback Options in the Geometric Brownian motion model, see subsection 4.6.8.2 on page 69. So with the same construction of the Maximums-process, we get for the randomized phantom estimator

$$\frac{\partial \mathbb{E}[H(S(T))]}{\partial b_1} = \mathbb{E}\left[e^{-r\cdot T} \cdot n \cdot T \cdot c_{b_1} \cdot \left(\max(M^+(T) - K, 0) - \max(M^-(T) - K, 0)\right)\right].$$

Table 5.14 shows the results with following input parameters: $T = 1, S(0) = 100, r = 0.01, a = 1, b_1 = 0.01, b_2 = 0.01, n = 252, K = 100, l = 0.01$ (For FD)

5.5 Compound Poisson model

5.5.1 Sensitivity w.r.t. b_1

$$\frac{\partial}{\partial b_1}\mathbb{E}\left[r^{-rT} H(S_{b_1}(T))\right]$$

	Estimator	Variance	computational time	WNV
MVDr	83,67	55,95	9,73	544,14
MVDe	83,71	2,10	21,09	44,34
FD	89,27	197,67	18,47	3650,58

Table 5.14: Sensitivity w.r.t. b_1 of a Lookback Option in the VG-model, $N = 200$.

In the case of a Compound Poisson model the log returns follow

$$\log S(i+1) - \log S(i) = \left(\sum_{k=1}^{N(\frac{1}{n})} (p \cdot V_1 - (1-p) \cdot V_2) \right)$$

with

$$V_1 \sim \text{Gamma}(a_1, b_1)$$
$$V_2 \sim \text{Gamma}(a_2, b_2).$$

Further

$$[\log S(i+1) - \log S(i)]^+ = \left(\sum_{k=1}^{N(\frac{1}{n})} \left(p \cdot V_1^+ - (1-p) \cdot V_2\right) \right)$$
$$[\log S(i+1) - \log S(i)]^- = \left(\sum_{k=1}^{N(\frac{1}{n})} (p \cdot V_1 - (1-p) \cdot V_2) \right)$$

with

$$V_1^+ \sim \text{Gamma}(a_1 + 1, b_1)$$

and

$$c_{b_1} = \frac{a_1}{b_1}.$$

Example (Sensitivity w.r.t. $b_1$for a Plain Vanilla Call Option)**.**

A simulation with input parameters: $r = 0.01, \lambda = 10, a_1 = 0.01, b_1 = 0.01, a_2 = 0.01, b_2 = 0.01, p = 0,5, \text{start} = 100, K = 100, n = 252, l = 0,01$ yields results according to Table 5.15.

5.5.2 Sensitivity w.r.t. λ

$$\frac{\partial}{\partial \lambda} \mathbb{E}\left[r^{-rT} H(S_\lambda(T))\right]$$

Considering the Compound Poisson process

	Estimator	Variance	computational time	WNV
MVDr	30,243	99,232	0,262	26,04
MVDe	30,679	4,982	0,943	4,70
FD	33,008	1027,01	0,242	248,73

Table 5.15: Sensitivity w.r.t. b_1 of a Vanilla Option in the CP-model, $N = 200$.

	Estimator	Variance	computational time	WNV
MVDr	0,054	0,0004	0,221	0,0007
FD	0,7503	293,46	0,294	86,375

Table 5.16: Sensitivity w.r.t. λ of a Vanilla Option in the CP-model, $N = 200$.

$$S(t) = S(0)e^{\sum_{k=1}^{N(t)} Z_k}$$

with $N(t) \sim \text{Poisson}(\lambda t)$, then for the weak derivative we have

$$\begin{aligned} S(t)^+ &= S(0)e^{\sum_{k=1}^{N(t)+1} Z_k} \\ S(t)^- &= S(0)e^{\sum_{k=1}^{N(t)} Z_k} \\ c_\lambda &= t. \end{aligned}$$

So for the postive phantom we have just to simulate one additional event.

Example (Sensitivity w.r.t. λ for a Vanilla Call Option)**.**

$$\frac{\partial \mathbb{E}\left(gH(S(T))\right)}{\partial \lambda} = \mathbb{E}\left[e^{-rT} \cdot c_\lambda \cdot \left(\max(S(T)^+ - K, 0) - \max(S(T)^- - K, 0)\right)\right]$$

A simulation with input parameters: $r = 0.01, \lambda = 10, a_1 = 0.01, b_1 = 0.01, a_2 = 0.01, b_2 = 0.01, p = 0,5, \text{start} = 100, K = 100, n = 252, l = 0,01$ yields results according to Table 5.16.

6 Appendices

Definition. A *σ-Algebra* $\mathcal{F}$ is a collection of sets $M \subset \Omega$ with

- $\emptyset, \Omega \in \mathcal{F}$
- If $M \in \mathcal{F} \Rightarrow M^c \in \mathcal{F}$
- If $(M_i)_{i=1}^{\infty} \in \mathcal{F} \Rightarrow M_i \in \mathcal{F}$

Let $X(t) : \Omega \to \mathbb{R}$ be a stochastic Process, then

Definition. A family $(\mathcal{F}_t \colon t \in T)$ of Sigma-Algebras is called *Filtration*, if for $s < t$ it holds that $\mathcal{F}_s \subseteq \mathcal{F}_t$. Furthermore if $X(t), t \in T$ is a stochastic process and $\mathcal{F}_t$ is the Sigmaalgebra which is generated of the random variables $X(s), s < t$, then we call $(\mathcal{F}_t, t \in T)$ the *natural Filtration* of X.

Definition. A set function $\mathbb{P}(.) : M \to \mathbb{R}$, with $M \in \Omega$ is called a probability measure if Ω together with an σ-Algebra $\mathcal{F}$ and a probability measure $\mathbb{P}$ forms a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $T \subset \mathbb{R}$. Commonly $T \in \mathbb{N}_0, \mathbb{R}_+$. A *stochastic process* on T is a familiy of random variables $(X(t), t \in T)$ with values in $M \subset \mathbb{R}^d$.

Definition. A stochastic process $(X(t), t \in T)$ is called a *Markov process*, if the futures behaviour (distribution) of the process is just dependent of the current state of the process relatively to the past. Formal for $s_1 < ... < s_n < s < t$

$$\mathbb{P}\left(X(t) \in B | X(s_1), ..., X(s_n), X(s)\right) = \mathbb{P}\left(X(t) \in B | X(s)\right)$$

Definition. A stochastic process $(X(n))_{n=0,1,...}$ with finite state space $\mathcal{Z} = \{z_1, ..., z_s\}$ is called *Markov Chain* (with discrete time), if for all $n \in \mathbb{N}$ and all $w_0, ..., w_n \in \mathcal{Z}$

Theorem.

$$\mathbb{P}(M(0) = w_0, M(1) = w_1, ..., M(n) = w_n) = \mathbb{P}(M(0) = w_0) \prod_{i=1}^{n} \mathbb{P}(M(i) = w_i | M(i-1) = w_{i-1})$$

[Pflug96].

Definition (Characteristic function)**.** the characteristic function is defined as the expected value of e^{isX}

$$\varphi_X(s) = \mathbb{E}\left(e^{isX}\right) = \int_{-\infty}^{\infty} e^{isx} dF_X(x)$$

- $X \sim Poisson(\lambda) \longrightarrow \varphi_X(s) = exp\left(\lambda(e^{is} - 1)\right)$
- $X \sim N(\mu, \sigma^2) \longrightarrow \varphi_X(s) = exp\left(is\mu - \frac{\sigma^2}{2}s^2\right)$
- $X \sim Gamma(a, b) \longrightarrow \varphi_X(s) = (1 - bis)^{-a} \quad b > 0$

The characteristic function of a distribution always exists, even when the probability density function or moment-generating function do not. Characteristic function can be used to find moments of a random variable. Provided that the n^{th} moment exists, the characteristic function can be differentiated n times and

$$E(X^n) = i^{-n}\varphi_X^{(n)}(0)$$

Definition. A random variable X is called infinitely divisible (or X has an infinitely divisible distribution), if for every $n \in \mathbb{N}$, there exist i.i.d. random variables $Z_1^{(1/n)}, ..., Z_n^{(1/n)}$ such that

$$X \overset{\mathcal{L}}{=} Z_1^{(1/n)} + ... + Z_n^{(1/n)}$$

Alternatively, we can characterize an infinitely divisible random variable by its characteristic function

Theorem. *X is infinitely divisible iff $\sqrt[n]{\varphi_X(u)}$ is a characteristic function for all n.*

Example. $X \sim N(\mu, \sigma^2)$:
$\varphi_X(t) = \exp\left[it\mu - \frac{\sigma^2}{2}t^2\right] \qquad \sqrt[n]{\varphi_X(u)} = exp\left[it\frac{\mu}{n} - \frac{\sigma^2}{2n}t^2\right]$

Example. $X \sim Poisson(\lambda)$:
$\varphi_X(t) = exp\left(\lambda(e^{it} - 1)\right) \qquad \sqrt[n]{\varphi_X(u)} = exp\left(\frac{\lambda}{n}(e^{it} - 1)\right)$

Example. $X \sim Gamma(a, b)$:
$\varphi_X(t) = (1 - bit)^{-a} \qquad \sqrt[n]{\varphi_X(u)} = (1 - bit)^{-\frac{a}{n}}$

Theorem (Bochner's theorem)**.** *φ is a characteristic function iff*

- *$\varphi$is nonnegative definite ($\sum_{i,j} \alpha_i\alpha_j\varphi(t_i - t_j) \geq 0$)*
- $\varphi(0) = 1$
- *φ is continuous at* 0

Definition. A random variable X is said to be stable (or X has a stable distribution), if it has the property that a positive linear combination of independent copies of X has the same distribution, up to location and scale parameters, i.e. X is stable if $\forall\, b_1, ..., b_d \in \mathbb{R}^+$, $\exists\, a$ and c such that

$$b_1X_1 + ... + b_dX_d = a \overset{\mathcal{L}}{+} cX$$

or in form of characteristic functions

$$\varphi(b_1t) \cdots \varphi(b_dt) = \varphi(ct)e^{iat}$$

see [6].

Remark. If $\varphi_X(u)$ is the characteristic function of a stable random variable X then the random variable is infinitely divisible and can be represented as a sum of i.i.d. random variables

$$X = Z_1 + Z_2 + ... + Z_n \stackrel{\mathcal{L}}{=} a_n + b_n Z_1$$

Example (Central Limit Theorem). $Z_1, ..., Z_n \sim N(0, \sigma^2)\, i.i.d.$
$X = Z_1 + Z_2 + ... + Z_n \stackrel{\mathcal{L}}{=} \sqrt{n} Z_1$

Definition (Signed measure). Given a measurable space (X, Σ), a signed measure γ is a function for which for all $A \in \Sigma$

- $A \to \gamma(A) \in \mathbb{R}$
- $\gamma(\oslash) = 0$
- $\gamma\left(\bigcup A_i\right) = \sum \gamma\left(A_i\right)$

for any disjoint sets $A_1, A_2, ... \in \Sigma$.

Definition (Jordan Hahn decomposition). Considering a measurable space (X, Σ), and a signed measure γ defined on the σ-algebra Σ, there exist a unique decomposition into two sets C, C^c in Σ such that

- $C \cup C^c = X$ and $C \cap C^c = \oslash$
- $\forall\, A \in \Sigma$ with $A \subseteq C$ it holds $\gamma(A) \geq 0$
- $\forall\, A \in \Sigma$ with $A \subseteq C^c$ it holds $\gamma(A) \leq 0$.

Therefore every signed measure γ can be written as a difference of two positive measures γ^+ and γ^- with

$$\gamma^+(A) = \gamma\left(A \cap C\right)$$
$$\gamma^-(A) = -\gamma\left(A \cap C^c\right)$$

$$\gamma(A) = \gamma^+(A) - \gamma^-(A).$$

The pair (γ^+, γ^-) is called the Jordan-Hahn decomposition of γ.

Lemma (Scheffé). *Suppose that X_n, X are integrable and that X_n converges to X in probability. If*

$$\limsup_n \int |X_n(\omega)|^p d\mathbb{P}(\omega) \leq \int |X(\omega)|^p d\mathbb{P}(\omega),$$

then X_n converges to X also in the L^p-sense.

Theorem 7. *Suppose that for $\theta \in \mathbb{R}$,*

1. *$\theta \to f_\theta(.|.)$ is a.s. differentiable (at the point θ) with derivative $\dot{f}_\theta(.|.) \in L^p$.*

2. $f_{\theta_2}(.|.) - f_{\theta_1}(.|.) = \int_{\theta_1}^{\theta^2} \dot{f}_\theta(.|.)d\theta,$ *a.s., for* $\theta_1 < \theta_2$

3. $\theta \mapsto \int |\dot{f}_\theta(\omega|.)|^p d\mu(\omega)$ *is continuous.*

Then $\dot{f}_\theta(.|.)$ *is the* L^p*-derivative of* $f_\theta(.|.)$*.*

Proof. By 2. and 3. and Hölder's inequality, for $s > 0\backslash$

$$\begin{aligned}
& \limsup_{s\downarrow 0} \int \frac{|f_{\theta+s}(\omega|.) - f_\theta(\omega|.)|^p}{s^p} d\mu(\omega) \\
= \; & \limsup_{s\downarrow 0} \frac{1}{s^p} \int |\int_\theta^{\theta+s} \dot{f}_v(\omega|.)dv|^p d\mu(\omega) \\
\leq \; & \limsup_{s\downarrow 0} \frac{1}{s^p} \int \int_\theta^{\theta+s} |\dot{f}_v(\omega|.)|^p dv \left[\int_\theta^{\theta+s} 1dv \right]^{p-1} d\mu(\omega) \\
= \; & \limsup_{s\downarrow 0} \frac{1}{s} \int_\theta^{\theta+s} \int |\dot{f}_v(\omega|.)|^p d\mu(\omega)dv = \int |\dot{f}_\theta(\omega|.)|^p d\mu(\omega).
\end{aligned}$$

By Scheffés Lemma this implies that $\frac{1}{s}\left[f_{\theta+s}(.|.) - f_\theta(.|.)\right]$ converges in L^p to $\dot{f}_\theta(.|.)$. The case for $s < 0$ is similar. $\square$

List of Figures

2.4.1 Brownian Motion with drift $\mu = 0.1$ and diffusion $\sigma = 0.5$. 17
2.4.2 Poisson process with $\lambda = 20$. 18
2.4.3 Compound Poisson process with $\lambda = 30, a_1 = 1, a_2 = 1, b_1 = 1, b_2 = 1$ and $p = 0.5$. 21
2.4.4 Gamma process with $a = 10$ and $b = 1$. 22
2.4.5 Variance-Gamma process with $a = 10, b = 2$ and $c = 3$. 24

4.6.1 MVDe, $n = 10$, $t = 1$. 60
4.6.2 MVD (random point), $n = 10$, $t = 1$, $\tau = 4$. 61
4.6.3 MVDc(2), $n = 10$, $t = 1$, $k = 2$, $m = 5$. 62
4.6.4 Example: Maximums-process. 70

5.1.1 Lookback Option: Example of generated phantoms. 76

List of Tables

2.1 List of Rainbow Options. 10

4.1 Distributions and their characteristic functions. 52
4.2 Examples of Weak derivatives 54

5.2 The Rho for a Digital Option in the BS-model, $N = 200$ 75
5.1 The Rho for a Vanilla Option in the BS-model, $N = 200$. 75
5.4 The Rho for a Lookback Option in the BS-model via the Maximums Process , $N = 200$. 77
5.3 The Rho for a Lookback Option in the BS-model , $N = 200$. 77
5.5 The Rho for an Asian Option in the BS-model , $N = 200$. 78
5.6 The Rho for an Asian Option in the BS-model via the Average process, $N = 200$. 79
5.7 The Vega for a Vanilla Option in the BS-model, $N = 200$. 80
5.8 The Delta for a Vanilla Option in the BS-model, $N = 200$. 81
5.9 The Delta for a Digital Option in the BS-model, $N = 200$. 82
5.10 Sensitivity w.r.t. λ of a Vanilla Option in the Poisson model, $N = 200$ . . 83
5.11 Sensitivity w.r.t. λ of a Digital Option in the Poisson model, $N = 200$ . . 83
5.12 Sensitivity w.r.t. b of a Vanilla Option in the Gamma model, $N = 200$ . . 84
5.13 Sensitivity w.r.t. b_1 of a Vanilla Option in the VG-model, $N = 200$. 85
5.14 Sensitivity w.r.t. b_1 of a Lookback Option in the VG-model, $N = 200$. . . 86
5.15 Sensitivity w.r.t. b_1 of a Vanilla Option in the CP-model, $N = 200$. 87
5.16 Sensitivity w.r.t. λ of a Vanilla Option in the CP-model, $N = 200$. 87

Bibliography

[1] David Applebaum. Lévy processes- from probability theory to finance and quantum groups. *Notices of the American Math Soc.*, 51:1336–1347, 2004.

[2] P.Carr B. Madan and C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[3] Georg Pflug Bernd Heidergott, Felisa J. Vázquez-Abad and T. Farenhorst-Yuan. Gradient estimation for discrete-event systems by measure-valued differentiation. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 20 Issue 1, 2010.

[4] B.Heidergott and F.J. Vázquez-Abad. Measure-valued differentiation for random horizon problems. *Markov Processes Related Fields*, 12:509–536, 2006.

[5] W. Volk-Makarewicz B.Heidergott, F.J. Vázquez-Abad. Sensitivity estimation for gaussian sys. *European Journal of Operational Research*, 187:193–207, 2008.

[6] Leo Breiman. *Probability*. Siam, 1992.

[7] John C.Hull. *Fundamentals of Futures and Options Markets*. Pearson Education, fifth edition edition, 2005.

[8] P.P. Carr D.B. Madan and E.C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[9] J. Lebuchoux P.-L.Lions E. Fournie, J.-M. Lasry and N. Touzi. Applications of malliavin calculus to monte carlo methods in finance. *Fin*, 3:391–412, 1999.

[10] D.Madan E.Eberlein and W.Schoutens. Capital requirements, the option surface, market, credit and liquidity risk. 2011.

[11] F. Proske G. Di Nunno, B. Oksendal. *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer, 2009.

[12] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2004.

[13] Paul Glasserman and Zongjian Liu. Estimating greeks in simulating lévy-driven models. 2008.

[14] P. Glynn and W. Whitt. The asymptotic efficiency of simulation estimators. *Operat*, 40:505–520, 1992.

[15] Karl Grill. Theorie stochastischer prozesse. Lecture Notes.

[16] A. Kohatsu-Higa H.-P. Bermin and M. Montero. Local vega index and variance reduction methods. *Mathematical Finance*, 13:85–97, 2003.

[17] Bernd Heidergott and Arie Hordijk. Single-run gradient estimation via measure-valued differentiation. *IEEE Transactions on Aut*, 49:1843–1846, 2004.

[18] Bernd Heidergott and Haralambie Leahu. Weak differentiability of product measures. *Mathematics of Operations Research*, 35:27–51, 2010.

[19] I.Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1991.

[20] Daniel J.Duffy. *Finite Difference Methods in Financial Eng.* John Wiley & Sons, 2006.

[21] L.Dai. Perturbation analysis via coupling. *IEEE Transactions on Automat*, 45:614–628, 2000.

[22] Yuh-Dauh Lyuu and Huei-Wen Teng. Unbiased and efficient greeks of financial options. *Finance and Stochastics*, 15:141–181, 2011.

[23] Paul Malliavin. *Stochastic Analysis*. Springer, 1997.

[24] M.Broadie and P.Glasserman. Estimating security price derivatives using simulation. *Management Science*, 42:269–285, 1996.

[25] M.Fu and J.Q.Hu. *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*. Kluwer A, 1997.

[26] Miquel Montero and Arturo Kohatsu-Higa. Malliavin calculus applied to finance. *Physica*, A 320:548–570, 2003.

[27] M.Reiman and A.Weiss. Sensitivity analysis for simulations via likelihood ratios. *Operations Research*, 37:830–844, 1989.

[28] N.Chen and P.Glasserman. Malliavin greeks without malliavin calculus. *Stochastic Processes and Applications*, 117:1689–1723, 2007.

[29] David Nualart. *The Malliavin Calculus and Related Topics*. Springer, 1995.

[30] Sidney I.Resnick O.E. Barndorff-Nielsen, Thomas Mikosch. *Lévy processes: Theory and Applications*. Birkhäuser, 2001.

[31] Bernt Oksendal. *Stochastic Differential Equations*. Springer, third edition edition, 1992.

[32] Antonis Papapantoleon. An introduction to lévy processes with applications in finance. Lecture Notes.

[33] Georg Pflug. *Optimization of Stochastic Models*. Kluwer Academic Publishers, 1996.

[34] Georg Pflug and Werner Römisch. *Modelling, Measuring and Managing Risk*. World Scientific Publishing Co. Pte. Ltd., 2007.

[35] Sebastian Raible. *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. PhD thesis, Albert-Ludwig-Universität Freiburg, 2000.

[36] G.Müller R.A.Maller and A.Szimayer. Ornstein-uhlenbeck processes and extensions. *Handbook of Financial Time Series*, pages 421–437, 2009.

[37] Marta Sanz-Solé. *Malliavin Calculus*. EPFL Press, 2005.

[38] Wim Schoutens. *Lévy Processes in Finance*. Wiley Series in Probability and Statistics, 2003.

[39] Stephan Sturm. Calculation of the greeks by malliavin calculus. Master's thesis, 2004.

[40] Felisa J. Vázquez-Abad and Bernd Heidergott. Theory and methods of stochastic optimization. Lecture Notes.

[41] F.J. Vázquez-Abad and B. Heidergott. Measure-valued differentiation for markov chains. *J Optim Theory Appl*, 136:187–209, 2008.

[42] Alan Whitley. *Pricing of European, Bermudan and American Options under the Exponential Variance Gamma Process*. PhD thesis, St Anne's College, University Oxford, 2009.

[43] M. Winkel. Lecture notes on lévy processes and finance.

Abstract

Monte Carlo simulation methods have become more and more important in the financial sector in the past years. In this work, we introduce a new simulation method for the estimation of the derivatives of prices of financial contracts w.r.t. certain distributional parameters, called the "Greeks". In particular, we assume that the underlying financial process is a Lévy-type process in discrete time.

Our method is based on the measure-valued differentiation (MVD) approach, which allows to represent derivatives as differences of two processes, called the phantoms. We discuss the applicability of MVD for different types of option payoffs in combination with different types of models of the underlying and provide a framework for the applicability of MVD for path-dependent payoff functions, as Lookback Options or Asian Options.

Kurzbeschreibung

Monte Carlo Simulationen haben im Bereich des Finanzsektors in den letzten Jahren immer mehr an Bedeutung gewonnen. Diese Arbeit stellt eine neue Simulationsmethode vor um Preissensitivitäten von Optionen bezüglich eines bestimmten Verteilungsparameters, sogenannte "Greeks", zu schätzen. Im Besonderen wird angenommen, dass der zugrundeliegende Prozess einem Lévy-Typ Prozess in diskreter Zeit entspricht.

Unsere Methode basiert auf dem Ansatz der Maßwertigen Ableitung (MVD), welcher uns erlaubt die Ableitung als Differenz zweier Prozesse, sogenannten Phantomen, darzustellen. Wir diskutieren die Anwendbarkeit der Maßwertigen Ableitung für verschiedene Arten von Optionen in Kombination mit unterschiedlichen Modellen des zugrundeliegenden Prozesses. Des Weiteren liefern wir den Rahmen für die Verwendbarkeit der Maßwertigen Ableitung für pfadabhängige Optionen, wie zum Beispiel Lookback Optionen oder Asiatische Optionen.

Dipl.-Ing. Philipp Thoma
1020 Wien, Franzensbrückenstraße 15/10
Tel.: 0650/2204958
Email: philipp.thoma@gmx.net

Philipp Thoma **Curriculum Vitae**

Personal Information:

Name	Dipl.-Ing. Philipp Thoma
Date of birth	25. November 1982
Nationality	Austria
City of birth	Zell am See

Education

Since 2009, October	PhD study at the **Department of Statistics and Operations Research at University of Vienna,** Vienna , Austria
2009, June	**Diploma examination with commendation**, graduated in degreed - Engineering (Dipl.-Ing.). Master Thesis in Game Theory: Combinatorial Auctions
2007 - 2008	Study abroad at the **City College of New York (CCNY)**, New York, USA
2002 – 2009	Diploma study of technical mathematic at **Vienna University of Technology**, Vienna, Austria Focus: Operations Research, Game Theory, Econometrics
1993-2001	Grammar School in Zell am See
1989 – 1993	Elementary school in Zell am See

Languages:

Mother Tongue	German
other Languages	English (fluent), French (basic)

Computer literacy:

Matlab, R, Maple, Latex/Lyx