Mathematical Modeling and Computation in Finance: With Exercises and Python and MATLAB Computer Codes

Book · December 2019

DOI: 10.1142/q0236

2 authors:

Cornelis W Oosterlee
Utrecht University

Lech Grzelak
Utrecht University

Cornelis W Oosterlee • Lech A Grzelak

Mathematical Modeling and Computation in Finance

World Scientific

Mathematical Modeling and Computation in Finance

This page intentionally left blank

Mathematical Modeling and Computation in Finance

with Exercises and Python and MATLAB computer codes

Cornelis W Oosterlee
Centrum Wiskunde & Informatica (CWI) &
Delft University of Technology, The Netherlands

Lech A Grzelak
Delft University of Technology, The Netherlands

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Europe Ltd.
57 Shelton Street, Covent Garden, London WC2H 9HE
Head office: 5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

Library of Congress Control Number: 2019950785

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

MATHEMATICAL MODELING AND COMPUTATION IN FINANCE
With Exercises and Python and MATLAB Computer Codes

ISBN 978-1-78634-794-7
ISBN 978-1-78634-805-0 (pbk)

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/Q0236#t=suppl

Desk Editor: Shreya Gopi

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Dedicated to Anasja, Wim, Mathijs, Wim en Agnes (Kees)
Dedicated to my mum, brother, Anna and my whole family (Lech)

Preface

This book is discussing the interplay of stochastics (applied probability theory) and numerical analysis in the field of quantitative finance. The contents will be useful for people working in the financial industry, for those aiming to work there one day, and for anyone interested in quantitative finance.

Stochastic processes, and stochastic differential equations of increasing complexity, are discussed for the various asset classes, reaching to the models that are in use at financial institutions. Only in exceptional cases, solutions to these stochastic differential equations are available in closed form.

The typical models in use at financial institutions have changed over time. Basically, each time when the behavior of participants in financial markets changes, the corresponding stochastic mathematical models describing the prices may change as well. Also financial regulation will play its role in such changes. In the book we therefore discuss a variety of models for stock prices, interest rates as well as foreign-exchange rates. A basic notion in such a diverse and varying field is *"don't fall in love with your favorite model"*.

Financial derivatives are products that are based on the performance of another, uncertain, underlying asset, like on stock, interest rate or FX prices. Next to the modeling of these products, they also have to be priced, and the risk related to selling these products needs to be assessed. Option valuation is also encountered in the financial industry during the calibration of the stochastic models of the asset prices (fitting the model parameters of the governing SDEs so that model and market values of options match), and also in risk management when dealing with counterparty credit risk.

Advanced risk management consists nowadays of taking into account the risk that a counterparty of a financial contract may default (CCR, Counterparty Credit Risk). Because of this risk, fair values of option prices are adjusted, by means of the so-called Valuation Adjustments. We will also discuss the Credit Valuation Adjustment (CVA) in the context of risk management and derive the governing equations.

Option values are governed by partial differential equations, however, they can also be defined as expectations that need to be computed in an efficient, accurate and robust way. We are particularly interested in stochastic volatility based models, with the well-known Heston model serving as the point of reference.

As the computational methods to value these financial derivatives, we present a Fourier-based pricing technique as well as the Monte Carlo pricing method. Whereas Fourier techniques are useful when pricing basic option contracts, like European options, within the calibration procedure, Monte Carlo methods are often used when more involved option contracts, or more involved asset price dynamics are being considered.

By gradually increasing the complexity of the stochastic models in the different chapters of the book, we aim to present the mathematical tools for defining appropriate models, as well as for the efficient pricing of European options. From the equity models in the first 10 chapters, we move to short-rate and market interest rate models. We cast these models for the interest rate into the Heath-Jarrow-Morton framework, show relations between the different models, and we explain a few interest rate products and their pricing as well.

It is sometimes useful to combine SDEs from different asset classes, like stock and interest rate, into a correlated set of SDEs, or, in other words, into a system of SDEs. We discuss the hybrid asset price models with a stochastic equity model and a stochastic interest rate model.

Summarizing, the reader may encounter a variety of stochastic models, numerical valuation techniques, computational aspects, financial products and risk management applications while reading this book. The aim is to help readers progress in the challenging field of computational finance.

The topics that are discussed are relevant for MSc and PhD students, academic researchers as well as for quants in the financial industry. We expect knowledge of applied probability theory (Brownian motion, Poisson process, martingales, Girsanov theorem, ...), partial differential equations (heat equation, boundary conditions), familiarity with iterative solution methods, like the Newton-Raphson method, and a basic notion of finance, assets, prices, options.

Acknowledgment

Here, we would like to acknowledge different people for their help in bringing this book project to a successful end.

First of all, we would like to thank our employers for their support, our groups at CWI — Center for Mathematics & Computer Science, in Amsterdam and at the Delft Institute of Applied Mathematics (DIAM), from the Delft University of Technology in the Netherlands. We thank our colleagues at the CWI, at DIAM, and at Rabobank for their friendliness. Particularly, Nada Mitrovic is acknowledged for all the help. Vital for us were the many fruitful discussions, cooperations and input from our group members, like from our PhD students, the post-docs and also the several guests in our groups. In particular, we thank our dear colleagues Peter Forsyth, Luis Ortiz Gracia, Mike Staunton, Carlos Vazquez, Andrea Pascucci, Yuying Li, and Karel in't Hout for their insight and the discussions.

Proofreading with detailed pointers and suggestions for improvements has been very valuable for us and for this we would like to thank in particular Natalia Borovykh, Tim Dijkstra, Clarissa Elli, Irfan Ilgin, Marko Iskra, Fabien Le Floc'h, Patrik Karlsson, Erik van Raaij, Sacha van Weeren, Felix Wolf and Thomas van der Zwaard.

We got inspired by the group's PhD and post-doctoral students, in alphabetic order, Kristoffer Andersson, Anastasia Borovykh, Ki Wai Chau, Bin Chen, Fei Cong, Fang Fang, Qian Feng, Andrea Fontanari, Xinzheng Huang, Shashi Jain, Prashant Kumar, Coen Leentvaar, Shuaiqiang Liu, Peiyao Luo, Marta Pou, Marjon Ruijter, Beatriz Salvador Mancho, Luis Souto, Anton van der Stoep, Maria Suarez, Bowen Zhang, Jing Zhao, and Hisham bin Zubair. We would like to thank Shreya Gopi, her team and Jane Sayers at World Scientific Publishing for the great cooperation.

Our gratefulness to our families cannot be described in words.

This page intentionally left blank

Using this Book

The book can be used as a textbook for MSc and PhD students in applied mathematics, quantitative finance or similar studies. We use the contents of the book ourselves for two courses at the university. One is called "Computational Finance", which is an MSc course in applied mathematics, in a track called Financial Engineering, where we discuss most of the first 10 chapters, and the other course is "Special Topics in Financial Engineering", where interest rate models and products but also the risk management in the form of counterparty credit risk and credit valuation adjustment are treated. At other universities, these courses are also called "Financial Engineering" or "Quantitative Finance".

Exercises are attached to each chapter, and the software used to get the numbers in the tables and the curves in the figures is available.

Below most tables and figures in the book there are MATLAB and Python icons

indicating that the corresponding MATLAB and Python computer codes are available. In the e-book version, clicking on the icons will lead, via a hyperlink, to the corresponding codes. All codes are also available on a special webpage.

On the webpage `www.QuantFinanceBook.com` the solutions to all odd-numbered exercises are available.
Also the Python and MATLAB computer codes can be found on the webpage.
Instructors can access the full set of solutions by registering at `https://www.worldscientific.com/worldscibooks/10.1142/q0236`

The computer codes come with a no warranty disclaimer.

No warranty disclaimer

The computer codes provided in this book are provided by the authors "as is", meant to support the theory and numerical results in this book. Any express or implied warranties are disclaimed. We do not guarantee that these codes are free of errors. They are not written according to some standards, and they are not intended for any general application.
Using the computer codes for other purposes than checking the solutions of the specific problems in the book is at the user's own risk and responsibility. The authors and publisher disclaim all liabilities for direct or consequential damages resulting from the use of the programs.

We wish you enjoyable reading!

Contents

CHAPTER 1

Basics about Stochastic Processes

In this chapter:

We introduce some basics about stochastic variables and stochastic processes, like probability density functions, expectations and variances. The *basics about stochastic processes* are presented in **Section 1.1**. Martingales and the *martingale property* are explained in **Section 1.2**. The stochastic *Itô integral* is discussed in quite some detail in **Section 1.3**.
These basic entities from probability theory are fundamental in financial mathematics.

Keywords: stochastic processes, stochastic integral Itô integral, martingales.

1.1 Stochastic variables

We first discuss some known results from probability theory, and start with some facts about stochastic variables and stochastic processes.

1.1.1 Density function, expectation, variance

A real-valued random variable X is often described by means of its cumulative distribution function (CDF),

$$F_X(x) := \mathbb{P}[X \leq x],$$

and its probability density function (PDF),

$$f_X(x) := \mathrm{d}F_X(x)/\mathrm{d}x.$$

Let X be a continuous real-valued random variable with PDF $f_X(x)$. The expected value of X, $\mathbb{E}[X]$, is defined as:

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)\mathrm{d}x = \int_{-\infty}^{+\infty} x \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}\mathrm{d}x = \int_{-\infty}^{+\infty} x\mathrm{d}F_X(x),$$

provided the integral $\int_{-\infty}^{+\infty} |x| f_X(x)\mathrm{d}x$ is finite.

The variance of X, $\mathbb{V}\mathrm{ar}[X]$ is defined as:

$$\mathbb{V}\mathrm{ar}[X] = \int_{-\infty}^{+\infty} (x - \mathbb{E}[X])^2 f_X(x)\mathrm{d}x,$$

provided the integral exists.

For a continuous random variable X and some constant $a \in \mathbb{R}$ the expectation of an indicator function is related to the CDF of X, as follows,

$$\mathbb{E}[\mathbb{1}_{X\le a}] = \int_{\mathbb{R}} \mathbb{1}_{x\le a} f_X(x)\mathrm{d}x = \int_{-\infty}^{a} f_X(x)\mathrm{d}x =: F_X(a),$$

where $F_X(\cdot)$ is the CDF of X, and where the notation $\mathbb{1}_{X\in\Omega}$ stands for the indicator function of the set Ω, defined as follows:

$$\mathbb{1}_{X\in\Omega} = \begin{cases} 1 & X \in \Omega, \\ 0 & X \notin \Omega. \end{cases} \tag{1.1}$$

Definition 1.1.1 (Survival probability) *The* survival probability *is directly linked to the CDF. If X is a random variable which denotes the lifetime, for example, in a population, then $\mathbb{P}[X \le x]$ indicates the probability of* not *reaching the age x. The survival probability, defined by*

$$\mathbb{P}[X > x] = 1 - \mathbb{P}[X \le x] = 1 - F_X(x),$$

then indicates the probability of surviving a lifetime of length x. ◀

A basic and well-known example of a random variable is the normally distributed random variable. A normally distributed stochastic variable X, with expectation μ and variance σ^2, is governed by the following probability distribution function:

$$F_{\mathcal{N}(\mu,\sigma^2)}(x) = \mathbb{P}[X \le x] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-(z-\mu)^2}{2\sigma^2}\right)\mathrm{d}z. \tag{1.2}$$

Variable X then is said to have an $\mathcal{N}(\mu,\sigma^2)$-normal distribution. The corresponding probability density function reads:

$$f_{\mathcal{N}(\mu,\sigma^2)}(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathcal{N}(\mu,\sigma^2)}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right). \tag{1.3}$$

Example 1.1.1 (Expectation, normally distributed random variable) Let us consider $X \sim \mathcal{N}(\mu, 1)$. By the definition of the expectation, we find:

$$\begin{aligned}\mathbb{E}[X] &= \int_{\mathbb{R}} x f_X(x)\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x \exp\left(-\frac{(x-\mu)^2}{2}\right)\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} (z+\mu)\exp\left(-\frac{z^2}{2}\right)\mathrm{d}z \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} z\exp\left(-\frac{z^2}{2}\right)\mathrm{d}z + \frac{\mu}{\sqrt{2\pi}}\int_{\mathbb{R}} \exp\left(-\frac{z^2}{2}\right)\mathrm{d}z,\end{aligned}$$

where we have used $z = x - \mu$. The function in the first integral is an odd function, i.e., $g(-x) = -g(x)$, so that the integral over $\mathbb{R}$ is equal to zero. The second integral is recognized as an integral over the PDF of $\mathcal{N}(0, 1)$, so it equals $\sqrt{2\pi}$. Therefore, the expectation of $\mathcal{N}(\mu, 1)$ is equal to μ, i.e., indeed $\mathbb{E}[X] = \mu$. ♦

1.1.2 Characteristic function

In computational finance, we typically work with the density function of certain stochastic variables, like for stock prices or interest rates, as it forms the basis for the computation of the variable's expectation or variance. For some basic stochastic variables the density function is known in closed-form (which is desirable), however, for quite a few relevant stochastic processes in finance we do not know the corresponding density. Interestingly, we will see, for example, in Chapters 5, 7 and 8 in this book, that for some of these processes we can derive expressions for other important functions that contain important information regarding expectations and other quantities. One of these functions is the so-called *characteristic function*, another is the *moment-generating function*. So, we will also introduce these functions here.

The *characteristic function* (ChF), $\phi_X(u)$ for $u \in \mathbb{R}$ of the random variable X, is the Fourier-Stieltjes transform of the cumulative distribution function $F_X(x)$, i.e., with i the imaginary unit,

$$\boxed{\phi_X(u) := \mathbb{E}\left[\mathrm{e}^{iuX}\right] = \int_{-\infty}^{+\infty} \mathrm{e}^{iux}\mathrm{d}F_X(x) = \int_{-\infty}^{+\infty} \mathrm{e}^{iux} f_X(x)\mathrm{d}x.} \tag{1.4}$$

A useful fact regarding $\phi_X(u)$ is that it uniquely determines the distribution function of X. Moreover, the moments of random variable X can also be derived by $\phi_X(u)$, as

$$\mathbb{E}\left[X^k\right] = \frac{1}{i^k}\frac{\mathrm{d}^k}{\mathrm{d}u^k}\phi_X(u)\Big|_{u=0},$$

with i again the imaginary unit, for $k \in \{0, 1, ...\}$, assuming $\mathbb{E}[|X|^k] < \infty$.

A relation between the characteristic function and the *moment generating function*, $\mathcal{M}_X(u)$, exists:

$$\mathbb{E}\big[X^k\big] = \frac{1}{i^k}\frac{\mathrm{d}^k}{\mathrm{d}u^k}\phi_X(u)\Big|_{u=0} \stackrel{\text{def}}{=} \frac{1}{i^k}\frac{\mathrm{d}^k}{\mathrm{d}u^k}\int_{-\infty}^{+\infty}\mathrm{e}^{iux}\mathrm{d}F_X(x)\Big|_{u=0} = \frac{\mathrm{d}^k}{\mathrm{d}u^k}\phi_X(-iu)\Big|_{u=0},$$

where the equation's right-hand side represents the moment-generating function, defined as

$$\mathcal{M}_X(u) := \phi_X(-iu) = \mathbb{E}\left[\mathrm{e}^{uX}\right]. \tag{1.5}$$

A relation exists between the moments of a positive random variable Y and the characteristic function for the log transformation $\phi_{\log Y}(u)$. For $X = \log Y$, the corresponding characteristic function reads:

$$\begin{aligned}\phi_{\log Y}(u) = \mathbb{E}\left[\mathrm{e}^{iu\log Y}\right] &= \int_0^\infty \mathrm{e}^{iu\log y} f_Y(y)\mathrm{d}y\\ &= \int_0^\infty y^{iu} f_Y(y)\mathrm{d}y.\end{aligned} \tag{1.6}$$

Note that we use $\log Y \equiv \log_e Y \equiv \ln Y$. By setting $u = -ik$, we have:

$$\phi_{\log Y}(-ik) = \int_0^\infty y^k f_Y(y)\mathrm{d}y \stackrel{\text{def}}{=} \mathbb{E}[Y^k]. \tag{1.7}$$

The derivations above hold for those variables for which the characteristic function for the log-transformed variable is available.

Example 1.1.2 (Density, characteristic function of normal distribution) In Figure 1.1 the CDF, PDF (left side picture) and the characteristic function (right side picture) of the normal distribution, $\mathcal{N}(10, 1)$, are displayed. The PDF and CDF are very smooth functions, whereas the characteristic function ("the Fourier transform of the density function") is an oscillatory function in the complex plane. ♦

Another useful function is the *cumulant characteristic function* $\zeta_X(u)$, defined as the logarithm of the characteristic function, $\phi_X(u)$:

$$\zeta_X(u) = \log \mathbb{E}\left[\mathrm{e}^{iuX}\right] = \log \phi_X(u).$$

The kth moment, $m_k(\cdot)$, and the kth cumulant, $\zeta_k(\cdot)$, can be determined by:

$$m_k(\cdot) = (-i)^k\frac{\mathrm{d}^k}{\mathrm{d}u^k}\phi_X(u)\Big|_{u=0}, \quad \zeta_k(\cdot) = (-i)^k\frac{\mathrm{d}^k}{\mathrm{d}u^k}\log\phi_X(u)\Big|_{u=0}, \tag{1.8}$$

where $(-i)^k \equiv i^{-k}$ for $k \in \mathbb{N}$, and with $\phi_X(u)$ defined in (1.4).

1.1.3 Cumulants and moments

Some properties follow directly from the definition of the characteristic function, such as

$$\phi_X(0) = 1, \quad \phi_X(-i) = \mathbb{E}\left[\mathrm{e}^X\right].$$

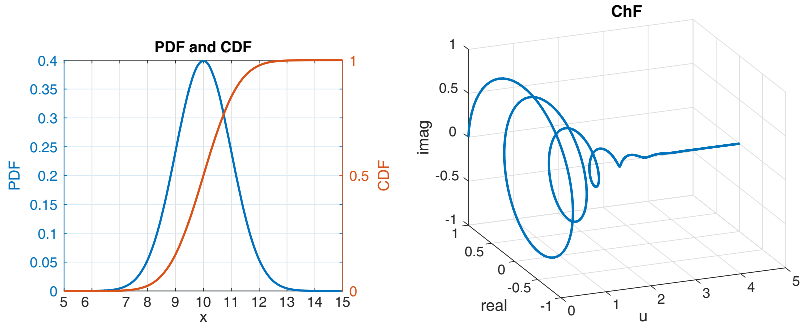


Figure 1.1: *The CDF, PDF and characteristic function for an* $\mathcal{N}(10,1)$ *random variable.*

By the definition of the moment-generating function, we also find

$$\mathcal{M}_X(u) \equiv \int_{\mathbb{R}} \mathrm{e}^{ux} f_X(x)\mathrm{d}x. \tag{1.9}$$

Assuming that all moments of X are finite, the moment-generating function (1.9) admits a MacLaurin series expansion:

$$\begin{aligned}\mathcal{M}_X(u) &= \int_{\mathbb{R}} \mathrm{e}^{ux} f_X(x)\mathrm{d}x \\ &= \sum_{k=0}^{\infty} \frac{u^k}{k!} \int_{\mathbb{R}} x^k f_X(x)\mathrm{d}x =: \sum_{k=0}^{\infty} m_k \frac{u^k}{k!},\end{aligned} \tag{1.10}$$

with the *raw moments*, $m_k = \int_{\mathbb{R}} x^k f_X(x)\mathrm{d}x$, for $k = 0, 1, \ldots$.
On the other hand, the cumulant-generating function is defined as:

$$\zeta_X(u) \equiv \log \mathcal{M}_X(u). \tag{1.11}$$

Using again the MacLaurin expansion gives us:

$$\zeta_X(u) = \sum_{k=0}^{\infty} \frac{\partial^k \zeta_X(u)}{\partial u^k}\bigg|_{u=0} \frac{u^k}{k!} =: \sum_{k=0}^{\infty} \zeta_k \frac{u^k}{k!}, \tag{1.12}$$

where $\zeta_k \equiv \frac{\partial^k \zeta_X(u)}{\partial u^k}\Big|_{u=0}$ is the k-th cumulant.

By (1.11) a recurrence relation between the raw moments and the cumulants is obtained, i.e.,

$$\mathcal{M}_X(u) = \sum_{k=0}^{\infty} m_k \frac{u^k}{k!} = \exp\left(\sum_{k=0}^{\infty} \zeta_k \frac{u^k}{k!}\right).$$

Using the results from (1.11), the first four derivatives of the cumulant-generating function can be expressed as,

$$\begin{aligned}
\frac{\mathrm{d}\zeta_X(t)}{\mathrm{d}t} &= \frac{1}{\phi_X(-it)}\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t},\\
\frac{\mathrm{d}^2\zeta_X(t)}{\mathrm{d}t^2} &= -\frac{1}{(\phi_X(-it))^2}\left(\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t}\right)^2 + \frac{1}{\phi_X(-it)}\frac{\mathrm{d}^2\phi_X(-it)}{\mathrm{d}t^2},\\
\frac{\mathrm{d}^3\zeta_X(t)}{\mathrm{d}t^3} &= \frac{2}{(\phi_X(-it))^3}\left(\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t}\right)^3 - \frac{3}{(\phi_X(-it))^2}\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t}\frac{\mathrm{d}^2\phi_X(-it)}{\mathrm{d}t^2}\\
&\quad + \frac{1}{\phi_X(-it)}\frac{\mathrm{d}^3\phi_X(-it)}{\mathrm{d}t^3},\\
\frac{\mathrm{d}^4\zeta_X(t)}{\mathrm{d}t^4} &= -\frac{6}{(\phi_X(-it))^4}\left(\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t}\right)^4 + \frac{12}{(\phi_X(-it))^3}\left(\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t}\right)^2\frac{\mathrm{d}^2\phi_X(-it)}{\mathrm{d}t^2}\\
&\quad - \frac{3}{(\phi_X(-it))^2}\left(\frac{\mathrm{d}^2\phi_X(-it)}{\mathrm{d}t^2}\right)^2 - \frac{4}{(\phi_X(-it))^2}\frac{\mathrm{d}\phi_X(-it)}{\mathrm{d}t}\frac{\mathrm{d}^3\phi_X(-it)}{\mathrm{d}t^3}\\
&\quad + \frac{1}{\phi_X(-it)}\frac{\mathrm{d}^4\phi_X(-it)}{\mathrm{d}t^4}.
\end{aligned}$$

With these derivatives of the cumulant-generating function, and the identity $\phi_X(0) = 1$, the first four cumulants are found as,

$$\begin{aligned}
\zeta_1 &= \frac{\mathrm{d}\zeta_X(t)}{\mathrm{d}t}|_{t=0}, \quad \zeta_2 = \frac{\mathrm{d}^2\zeta_X(t)}{\mathrm{d}t^2}|_{t=0},\\
\zeta_3 &= \frac{\mathrm{d}^3\zeta_X(t)}{\mathrm{d}t^3}|_{t=0}, \quad \zeta_4 = \frac{\mathrm{d}^4\zeta_X(t)}{\mathrm{d}t^4}|_{t=0}.
\end{aligned}$$

For a random variable X, with μ its mean X, σ^2 its variance, γ_3 the skewness and γ_4 the kurtosis by the relation between the cumulants and moments, the following equalities can be found,

$$\begin{aligned}
\zeta_1 &= m_1 = \mu,\\
\zeta_2 &= m_2 - m_1^2 = \sigma^2,\\
\zeta_3 &= 2m_1^3 - 3m_1m_2 + m_3 = \gamma_3\sigma^3,\\
\zeta_4 &= -6m_1^4 + 12m_1^2m_2 - 3m_2^2 - 4m_1m_3 + m4 = \gamma_4\sigma^4, \qquad (1.13)
\end{aligned}$$

Recall that skewness is a measure of asymmetry around the mean of a distribution, whereas kurtosis is a measure for the tailedness of a distribution.

When the moments or the cumulants are available, the associated density can, at least formally, be recovered. Based on Equation (1.4), the probability density function can be written as the inverse Fourier transform of the characteristic function:

$$\boxed{f_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \phi_X(u)\mathrm{e}^{-iux}\mathrm{d}u.} \tag{1.14}$$

By the definition of the characteristic function and the MacLaurin series expansion of the exponent around zero, we find:

$$\phi_X(u) = \mathbb{E}\left[\mathrm{e}^{iuX}\right] = \sum_{k=0}^{\infty} \frac{(iu)^k}{k!}\mathbb{E}[X^k] = \sum_{k=0}^{\infty} \frac{(iu)^k}{k!} m_k, \tag{1.15}$$

with m_k as in (1.10). Thus, Equation (1.14) equals:

$$f_X(x) = \frac{1}{2\pi}\sum_{k=0}^{\infty}\frac{m_k}{k!}\int_{\mathbb{R}}(iu)^k\mathrm{e}^{-iux}\mathrm{d}u. \tag{1.16}$$

Recall the definition of the *Dirac delta function and its k-th derivative*, as follows,

$$\delta(x) = \begin{cases} +\infty, & x = 0, \\ 0, & x \neq 0. \end{cases} \tag{1.17}$$

and

$$\delta(x) = \frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{e}^{-iux}\mathrm{d}u, \quad \delta^{(k)}(x) = \frac{\mathrm{d}^k}{\mathrm{d}x^k}\delta(x) = \frac{1}{2\pi}\int_{\mathbb{R}}(-iu)^k\mathrm{e}^{-iux}\mathrm{d}u.$$

Now, that Equation (1.16) is also written as:

$$f_X(x) = \sum_{k=0}^{\infty}(-1)^k\frac{m_k}{k!}\delta^{(k)}(x). \tag{1.18}$$

Two-dimensional densities

The joint CDF of two random variables, X and Y, is the function $F_{X,Y}(\cdot,\cdot) : \mathbb{R}^2 \to [0,1]$, which is defined by:

$$F_{X,Y}(x,y) = \mathbb{P}[X \leq x, Y \leq y].$$

If X and Y are continuous variables, then the *joint PDF of X and Y* is a function $f_{X,Y}(\cdot,\cdot) : \mathbb{R}^2 \to \mathbb{R}^+ \cup \{0\}$, such that:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y},$$

For any event A, it follows that

$$\mathbb{P}[(X,Y) \in A] = \iint_A f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

We will also use the vector notation in this book, particularly from Chapter 7 on, where we will then write $\mathbf{X} = [X,Y]^{\mathrm{T}}$ and $F_{\mathbf{X}},\ f_{\mathbf{X}}$, respectively.

As the joint PDF is a true probability function, we have $f_{X,Y}(x,y) \geq 0$, for any $x, y \in \mathbb{R}$ and

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 1.$$

Given the joint distribution of (X,Y), the expectation of a function $h(X,Y)$ is calculated as:

$$\mathbb{E}[h(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x,y)f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y.$$

The *conditional PDF of Y, given $X = x$,* is defined as:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad -\infty < y < \infty.$$

Moreover, the *conditional expectation of X, given $Y = y$,* is defined as the mean of the conditional PDF of X, given $Y = y$, i.e.,

$$\mathbb{E}[X|Y=y] = \int_{-\infty}^{+\infty} x f_{X|Y}(x|y)\mathrm{d}x = \int_{-\infty}^{+\infty} x\frac{f_{X,Y}(x,y)}{f_Y(y)}\mathrm{d}y.$$

Based on the joint PDF, we can also determine the *marginal densities*, as follows,

$$\begin{aligned} f_X(x) &= \frac{\mathrm{d}}{\mathrm{d}x}F_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}[X \leq x] \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\int_{-\infty}^{x}\left(\int_{-\infty}^{+\infty} f_{X,Y}(u,y)\mathrm{d}y\right)\mathrm{d}u \\ &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}y, \end{aligned}$$

and similarly for $f_Y(y)$.

Example 1.1.3 (Bivariate normal density functions) In this example, we show three bivariate normal density and distribution functions, with $\mathbf{X} = [X,Y]^{\mathrm{T}}$, and

$$\mathbf{X} \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1, & \rho\\ \rho, & 1\end{bmatrix}\right),$$

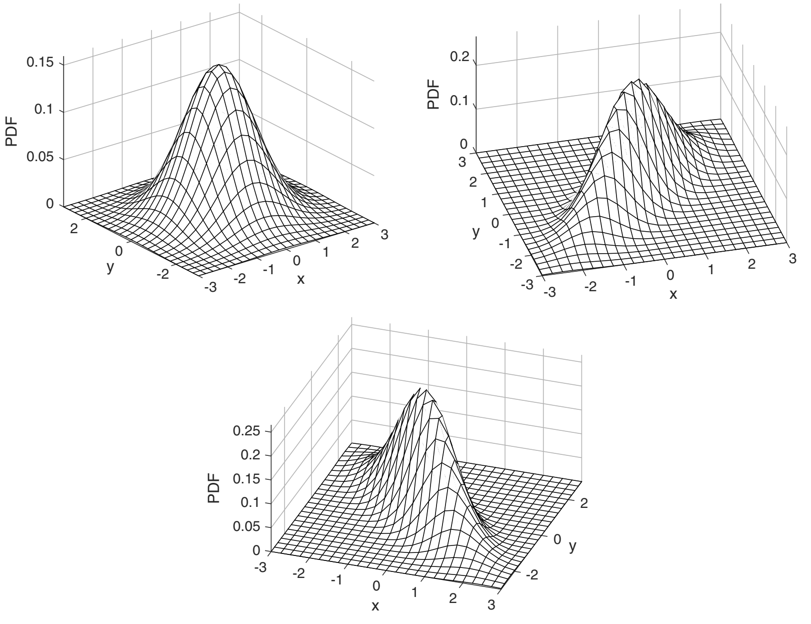


Figure 1.2: *Examples of two-dimensional normal probability density functions, with* $\rho = 0$ *(first),* $\rho = 0.8$ *(second) and* $\rho = -0.8$ *(third).*

in which the correlation coefficient is varied. Figure 1.2 displays the functions for $\rho = 0$ (first row), $\rho = 0.8$ (second row) and $\rho = -0.8$ (third row). Clearly, the correlation coefficient has an impact of the *direction* in these functions. ♦

1.2 Stochastic processes, martingale property

We will often work with stochastic processes for the financial asset prices, and give some basic definitions for them here.

A stochastic process, $X(t)$, is a collection of random variables indexed by a *time* variable t.

Suppose we have a set of calendar dates/days, $T_1, T_2, \ldots, T_m$. Up to *today*, we have observed certain state values of the stochastic process $X(t)$, see Figure 1.3. The past is known, and we therefore "see" the historical asset path. For the future we do not know the precise path but we may simulate the future according to some asset price distribution.

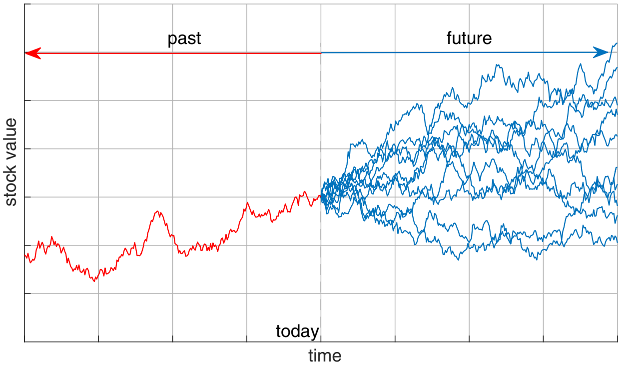


Figure 1.3: *Past and present in an asset price setting. We do not know the precise future asset path but we may simulate it according to some price distribution.*

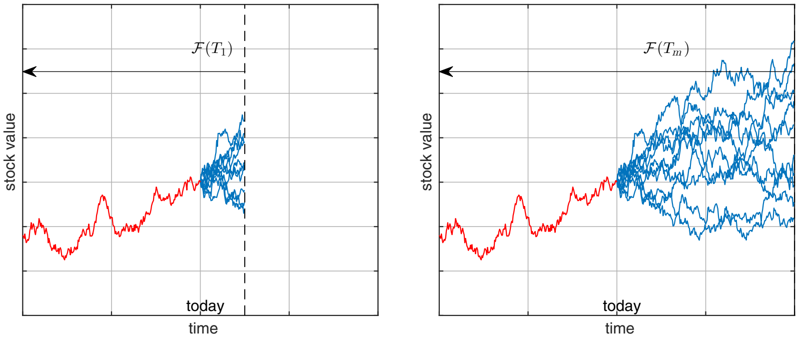


Figure 1.4: *Filtration figure, with $\mathcal{F}(t_0) \subseteq \mathcal{F}(T_1) \subseteq \mathcal{F}(T_2) \ldots \subseteq \mathcal{F}(T_m)$. When $X(t)$ is $\mathcal{F}(t_0)$ measurable this implies that at time t_0 the value of $X(t)$ is known. $X(T_1)$ is $\mathcal{F}(T_1)$ measurable, but $X(T_1)$ is a "future realization" which is not yet known at time t_0 ("today") and thus not $\mathcal{F}(t_0)$ measurable.*

The mathematical tool which helps us describe the knowledge of a stochastic process up-to a certain time T_i is the *sigma-field*, also known as sigma-algebra. The ordered sequence of sigma-fields is called a *filtration*, $\mathcal{F}(T_i) := \sigma(X(T_j) : 1 \leq j \leq i)$, generated by the sequence $X(T_j)$ for $1 \leq j \leq i$. The information available at time T_i is thus described by a filtration, see also Figure 1.4. As we consider a sequence of observation dates, $T_1, \ldots, T_i$, we deal in fact with a sequence of filtrations, $\mathcal{F}(T_1) \subseteq \cdots \subseteq \mathcal{F}(T_i)$.

If we write that a process is $\mathcal{F}(T)$-measurable, we mean that at any time $t \leq T$, the realizations of this process are known. A simple example for this may be the market price of a stock and its historical values, i.e., we know the stock values up to today exactly, but we do not know any future values. We then say "the stock is today measurable". However, when we deal with an SDE model for the stock price, the value may be T measurable, as we know the distribution for the period T of a financial contract.

A stochastic process $X(t)$, $t \geq 0$, is said to be adapted to the filtration $\mathcal{F}(t)$, if

$$\sigma(X(t)) \subseteq \mathcal{F}(t).$$

By the term "adapted process" we mean that a stochastic process "cannot look into the future". In other words, for a stochastic process $X(t)$ its realizations (paths), $X(s)$ for $0 \leq s < t$, are known at time s *but not yet at time* t.

1.2.1 Wiener process

Definition 1.2.1 (Wiener process) *A fundamental stochastic process, which is also commonly used in the construction of stochastic differential equations (SDEs) to describe asset price movements, is the Wiener process, also called Brownian motion. Mathematically, a Wiener process, $W(t)$, is characterized by the following properties:*

a. $W(t_0) = 0$, *(technically:* $\mathbb{P}[W(t_0) = 0] = 1$*),*

b. $W(t)$ *is almost surely*[a] *continuous,*

c. $W(t)$ *has independent increments, i.e.* $\forall\ t_1 \leq t_2 \leq t_3 \leq t_4$, $W(t_2) - W(t_1) \perp\!\!\!\perp W(t_4) - W(t_3)$, *with distribution* $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ *for* $0 = t_0 \leq s < t$, *i.e. the normal distribution with mean 0 and variance* $t - s$. ◀

[a] Almost surely convergence means that for a sequence of random variables X_m, the following holds: $\mathbb{P}\left[\lim_{m\to\infty} X_m = X\right] = 1$.

Example 1.2.1 Examples of processes that are adapted to the filtration $\mathcal{F}(t)$ are:

- $W(t)$ and $W^2(t) - t$, with $W(t)$ a Wiener process.
- $\max\limits_{0\leq s\leq t} W(s)$ and $\max\limits_{0\leq s\leq t} W^2(s)$.

Examples of processes that are not adapted to the filtration $\mathcal{F}(t)$ are:

- $W(t+1)$,
- $W(t) + W(T)$ for some $T > t$. ♦

1.2.2 Martingales

An important notion when dealing with stochastic processes is the martingale property.

Definition 1.2.2 (Martingale) *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where Ω is the set of all possible outcomes, $\mathcal{F}(t)$ is the sigma-field, and $\mathbb{Q}$ is a probability measure. A right continuous process $X(t)$ with left limits (so-called cádlág[a] process) for $t \in [0, T]$, is said to be a martingale with respect to the filtration $\mathcal{F}(t)$ under measure $\mathbb{Q}$, if for all $t < \infty$, the following holds:*

$$\mathbb{E}[|X(t)|] < \infty,$$

and

$$\mathbb{E}[X(t)|\mathcal{F}(s)] = X(s), \quad \text{with} \quad s < t,$$

where $\mathbb{E}[\cdot|\mathcal{F}]$ is the conditional expectation operator under measure $\mathbb{Q}$. ◀

[a]"continue à droite, limite à gauche" (French for right continuous with left limits).

The definition implies that the best prediction of the expectation of a martingale's future value is its present value, and

$$\begin{aligned}\mathbb{E}\left[X(t+\Delta t) - X(t)|\mathcal{F}(t)\right] &= \mathbb{E}[X(t+\Delta t)|\mathcal{F}(t)] - \mathbb{E}[X(t)|\mathcal{F}(t)]\\ &= X(t) - X(t) = 0,\end{aligned} \tag{1.19}$$

for some time interval $\Delta t > 0$.

Proposition 1.2.1 *The Wiener process $W(t), t \in [0, T]$ is a martingale.* ◀

Proof We check the martingale properties. First of all, $\mathbb{E}[|W(t)|] < \infty$, since

$$\mathbb{E}\left[|W(t)|\right] = \int_{-\infty}^{+\infty} |x| \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{x^2}{2t}} \mathrm{d}x = \frac{2}{\sqrt{2\pi t}} \int_0^{+\infty} x\mathrm{e}^{-\frac{x^2}{2t}} \mathrm{d}x.$$

By setting $z = \frac{x^2}{2t}$, so that $t\mathrm{d}z = x\mathrm{d}x$, we find

$$\mathbb{E}\left[|W(t)|\right] = \frac{2t}{\sqrt{2\pi t}} \int_0^{+\infty} \mathrm{e}^{-z}\mathrm{d}z = \frac{2t}{\sqrt{2\pi t}} < \infty \ \text{ for finite } \ t.$$

For Wiener process $W(t), t \in [0, T]$, we also find, using (1.19), that

$$\begin{aligned}\mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[W(s) + [W(t) - W(s)]|\mathcal{F}(s)]\\ &= \mathbb{E}[W(s)|\mathcal{F}(s)] + \mathbb{E}[W(t) - W(s)|\mathcal{F}(s)]\\ &= W(s) + 0 = W(s), \ \forall\, s, t > 0.\end{aligned}$$

■

1.2.3 Iterated expectations (Tower property)

Another important and useful concept is the concept of iterated expectations. The law of iterated expectations, also called the *tower property*, states that for any given random variable $X \in L^2$ (where L^2 indicates a so-called Hilbert space for which $\mathbb{E}[X^2(t)] < \infty$), which is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, and for any sigma-field $\mathcal{G} \subseteq \mathcal{F}$, the following equality holds:

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}], \quad \text{for} \quad \mathcal{G} \subseteq \mathcal{F}.$$

If we consider another random variable Y, which is defined on the sigma-field $\mathcal{G}$, so that $\mathcal{G} \subseteq \mathcal{F}$, then the above equality can be written as

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]], \quad \text{for} \quad \sigma(Y) \subseteq \sigma(X).$$

Assuming that both random variables, X and Y, are continuous on $\mathbb{R}$ and are defined on the same sigma-field, we can prove the equality given above, as follows

$$\begin{aligned}\mathbb{E}[\mathbb{E}[Y|X]] &= \int_{\mathbb{R}} \mathbb{E}[Y|X=x] f_X(x)\mathrm{d}x \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} y f_{Y|X}(y|x)\mathrm{d}y \right) f_X(x)\mathrm{d}x.\end{aligned}$$

By the definition of the conditional density, i.e. $f_{Y|X}(y|x) = f_{Y,X}(y,x)/f_X(x)$, we have:

$$\begin{aligned}\mathbb{E}[\mathbb{E}[Y|X]] &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} y \frac{f_{Y,X}(y,x)}{f_X(x)}\mathrm{d}y \right) f_X(x)\mathrm{d}x \\ &= \int_{\mathbb{R}} y \left(\int_{\mathbb{R}} f_{Y,X}(y,x)\mathrm{d}x \right) \mathrm{d}y \qquad (1.20) \\ &= \int_{\mathbb{R}} y f_Y(y)\mathrm{d}y \overset{\text{def}}{=} \mathbb{E}[Y].\end{aligned}$$

The conditional expectation will be convenient when dealing with continuous distributions. Let us take two independent random variables X and Y. Using the conditional expectation we can show the following equality:

$$\mathbb{P}[X < Y] = \int_{\mathbb{R}} \mathbb{P}[X < y] f_Y(y)\mathrm{d}y = \int_{\mathbb{R}} F_X(y) f_Y(y)\mathrm{d}y,$$

which can be proven as follows,

$$\begin{aligned}\mathbb{P}[X < Y] &= \mathbb{E}[\mathbb{1}_{X<Y}] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{X<Y}|Y=y]] = \int_{\mathbb{R}} \mathbb{P}[X < y|Y=y] f_Y(y)\mathrm{d}y.\end{aligned}$$

Since X and Y are independent variables, we have:

$$\mathbb{P}[X < Y] = \int_{\mathbb{R}} \mathbb{P}[X < y] f_Y(y)\mathrm{d}y = \int_{\mathbb{R}} F_X(y) f_Y(y)\mathrm{d}y.$$

The result given above can be used to show an example of a so-called *convolution*, which, for constant $c \in \mathbb{R}$ and two independent random variables X and Y, is defined as:

$$\mathbb{P}[X + Y < c] = \int_{\mathbb{R}} F_Y(c - x) f_X(x) \mathrm{d}x = \int_{\mathbb{R}} F_X(c - y) f_Y(y) \mathrm{d}y.$$

These integrals can be recognized as two expectations, $\mathbb{E}[F_Y(c-X)]$ and $\mathbb{E}[F_X(c-Y)]$.

1.3 Stochastic integration, Itô integral

For any differentiable function $\xi(t)$, one can use the following relation:

$$\int_0^T g(t) \mathrm{d}\xi(t) = \int_0^T g(t) \left(\frac{\mathrm{d}\xi(t)}{\mathrm{d}t} \right) \mathrm{d}t. \tag{1.21}$$

However, when $\xi(t)$ is a Wiener process, i.e. $\xi(t) \equiv W(t)$, Equality (1.21) is not valid, as the Brownian motion is nowhere differentiable.

Riemann-Stieltjes integration cannot be used when the integrand is based on a Wiener process. *Stochastic integration* can however be applied with the calculus, which is developed by the Japanese mathematician Kiyoshi Itô (1915–2008).

We consider the following stochastic differential equation,

$$\mathrm{d}I(t) = g(t) \mathrm{d}W(t), \quad \text{for } t \geq 0. \tag{1.22}$$

which is equivalent to the following *Itô integral*:

$$I(T) = \int_0^T g(t) \mathrm{d}W(t), \quad \text{for } T \geq 0, \tag{1.23}$$

where, in a certain interval $[0, T]$, the function $g(t)$ may represent a stochastic process, $g(t) := g(t, \omega)$. The variable ω then represents *randomness*, i.e. $\omega \in \Omega$ given the filtration $\mathcal{F}(t)$. The function $g(t)$ needs to satisfy the following two conditions:

1. $g(t)$ is $\mathcal{F}(t)$-measurable for any time t (in other words, the "process" $g(t)$ is an adapted process).

2. $g(t)$ is square-integrable, i.e.: $\mathbb{E}\left[\int_0^T g^2(t) \mathrm{d}t\right] < \infty$, $\forall\, T \geq 0$.

1.3.1 Elementary processes

For a given partition, $0 = t_0 < t_1 < \cdots < t_m = T$, of the time interval $[0, T]$, we make use of *elementary processes*, $\{g_m(t)\}_{m=0}^{\infty}$, with $g_m(t)$ a piecewise constant function. With the help of these elementary processes, we can formally define the Itô integral, as follows.

Definition 1.3.1 *For any square-integrable adapted process $g(t) = g(t, \omega)$, with continuous sample paths, the Itô integral is given by:*

$$I(T) \stackrel{\text{def}}{=} \int_0^T g(t)\mathrm{d}W(t) = \lim_{m\to\infty} I_m(T), \quad in \quad L^2. \tag{1.24}$$

Here, $I_m(T) = \int_0^T g_m(t)\mathrm{d}W(t)$ for some elementary process $\{g_m(t)\}_{m=0}^{\infty}$, satisfying:

$$\lim_{m\to\infty} \mathbb{E}\left[\int_0^T (g_m(t) - g(t))^2\, \mathrm{d}t\right] = 0. \tag{1.25}$$

◄

The existence of a sequence of elementary processes is presented with the help of the following theorems.

Theorem 1.3.1 (Dominated Convergence Theorem in L^p)
Let $\{\xi_m\}_{m\in\mathbb{N}}$ be a sequence of functions in L^p, $p > 0$, such that there exists a real-valued function $\bar{\xi} \in L^p$ with $|\xi_m| < \bar{\xi}$ for all $m \in \mathbb{N}$. Assume that $\{\xi_m\}_{m\in\mathbb{N}} \to \xi$, in a pointwise fashion. Then

$$||\xi_m - \xi||_{L^p} = \lim_{m\to\infty} \left(\mathbb{E}\left[|\xi_m - \xi|^p\right]\right)^{\frac{1}{p}} = 0.$$

The proof of this theorem can be found in stochastic calculus textbooks. We will use $p = 1$ and $p = 2$ below.

Theorem 1.3.2 *A sequence of elementary processes $\{g_m(t)\}_{m=0}^{\infty}$ exists, such that:*

$$\lim_{m\to\infty} \mathbb{E}\left[\int_0^T |g_m(t) - g(t)|^2 \mathrm{d}t\right] = 0. \tag{1.26}$$

Let us, for simplicity, assume here $T \in \mathbb{N}$. The objective of a proof of Theorem 1.3.2 is to find a sequence of elementary processes, $g_1(t), g_2(t), \ldots$, such that Equation (1.26) holds. To achieve this, we define the following elementary processes:

$$g_m(t) = \begin{cases} m\displaystyle\int_{\frac{k-1}{m}}^{\frac{k}{m}} g(s)\mathrm{d}s & \text{if} \quad t \in \left[\frac{k-1}{m}, \frac{k}{m}\right) \quad \text{for} \quad k = 1, 2, \ldots, mT, \\ 0 & \text{otherwise}, \end{cases} \tag{1.27}$$

The construction implies that $g_m(t)$ is in essence a step function, i.e. it is constant on each interval $t \in \left[\frac{k-1}{m}, \frac{k}{m}\right)$. However, because $g(t)$ is stochastic, each path of $g(t)$ yields a different constant realization of $g_m(t)$. Commonly, the function

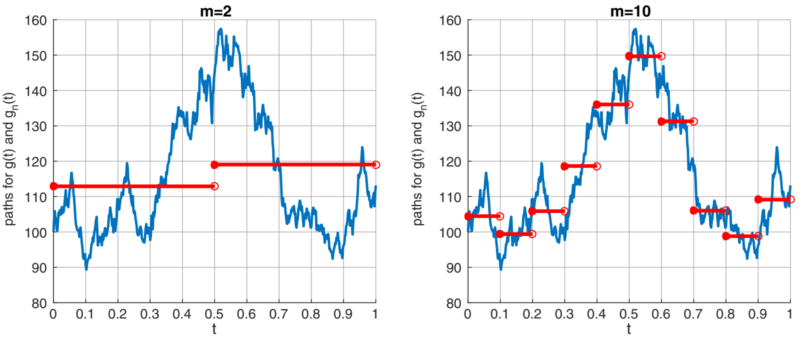


Figure 1.5: *Random step functions, approximating a stochastic function $g(t)$, with $m = 2$ and $m = 10$, respectively.*

in (1.27) is therefore called a *random step function*. We present an example in Figure 1.5.

By the Cauchy-Schwartz inequality, we obtain the following relation:

$$\begin{aligned}\int_{\frac{k-1}{m}}^{\frac{k}{m}} |g_m(t)|^2 \mathrm{d}t &= \int_{\frac{k-1}{m}}^{\frac{k}{m}} \left| m \int_{\frac{k-1}{m}}^{\frac{k}{m}} g(z)\mathrm{d}z \right|^2 \mathrm{d}t = \left(\frac{k}{m} - \frac{k-1}{m}\right) \left| m \int_{\frac{k-1}{m}}^{\frac{k}{m}} g(z)\mathrm{d}z \right|^2 \\ &\leq m\left(\frac{k}{m} - \frac{k-1}{m}\right) \int_{\frac{k-1}{m}}^{\frac{k}{m}} g^2(z)\mathrm{d}z = \int_{\frac{k-1}{m}}^{\frac{k}{m}} g^2(z)\mathrm{d}z \quad a.s. \qquad (1.28)\end{aligned}$$

Where the abbreviation "a.s." stands for the *almost surely*, which essentially means that this inequality needs to hold for any realization $g(t)$.

We now move on to the main part of the proof, i.e., showing the equality in (1.26).

By the assumption of the a.s. continuity sample paths of $g(t)$, we have:

$$\lim_{m\to\infty} \int_0^T |g_m(t) - g(t)|^2 \mathrm{d}t = 0 \;\; a.s. \qquad (1.29)$$

Defining $\xi_m := \int_0^T |g_m(t) - g(t)|^2 \mathrm{d}t$, we thus have $\lim_{m\to\infty} \xi_m = 0$.

Based on the inequality $(a + b)^2 \leq 2(a^2 + b^2)$, the following holds true:

$$\xi_m \equiv \int_0^T |g_m(t) - g(t)|^2 \mathrm{d}t \leq 2\int_0^T |g_m(t)|^2 \mathrm{d}t + 2\int_0^T |g(t)|^2 \mathrm{d}t.$$

We then find, using the inequality in (1.28),

$$\xi_m \leq 4\int_0^T |g(t)|^2 \mathrm{d}t =: \bar{\xi}. \qquad (1.30)$$

Since $\xi_m \to 0$ a.s. when $m \to \infty$ (1.29), $|\xi_m| < \bar{\xi}$ a.s. (1.30), and $\mathbb{E}[\bar{\xi}] < \infty$ by Theorem 1.3.1, we have:

$$||\xi_m - \xi||_{L^1} = \lim_{m\to\infty} \mathbb{E}\left[(\xi_m - 0)^1\right] = \lim_{m\to\infty} \mathbb{E}\left[\int_0^T |g_m(t) - g(t)|^2 \mathrm{d}t\right] = 0.$$

1.3.2 Itô isometry

Let's look at the discrete version of the Itô integral in some more detail:

$$I(T) := \int_0^T g(t)\mathrm{d}W(t) = \lim_{m\to\infty} \sum_{i=0}^{m-1} g(t_i)(W(t_{i+1}) - W(t_i)), \tag{1.31}$$

with $t_i = i\frac{T}{m}$. A proof of existence of the limit at the right-hand-side of (1.31) can be found in Theorem 1.3.3; more details and information regarding uniqueness can be found in the standard literature, see, for example, [Shreve, 2004].

The particular choice for the evaluation of $g(t)$ at the left-hand point of $[t_i, t_{i+1})$ is specific to Itô's calculus. If one evaluates the function at the mid-point, i.e., using $g((t_{i+1} + t_i)/2)$, integration is according to Stratonovich. Itô's integration has a preference in finance, as the left-hand time point indicates present time, whereas a stock price at the mid-point would rely on time points *in the future*.

We can make use of the property $\mathbb{E}[I(T)] \equiv \mathbb{E}[I(T)|\mathcal{F}(t_0)] = 0$ (a property we will encounter frequently in this chapter), which can be seen from the following derivation,

$$\begin{aligned}
\mathbb{E}[I(T)] &= \mathbb{E}\left[\lim_{m\to\infty} \sum_{i=0}^{m-1} g(t_i)\left(W(t_{i+1}) - W(t_i)\right)\right] \\
&= \lim_{m\to\infty} \sum_{i=0}^{m-1} \mathbb{E}\left[g(t_i)\left(W(t_{i+1}) - W(t_i)\right)\right] \\
&= \lim_{m\to\infty} \sum_{i=0}^{m-1} \mathbb{E}\left[g(t_i)|\mathcal{F}(t_0)\right]\mathbb{E}\left[W(t_{i+1}) - W(t_i)\right]. \qquad (1.32)
\end{aligned}$$

Since increments of a Brownian motion are independent with respect to stochastic variables and functions up to time t_i, and since the increments $W(t_{i+1}) - W(t_i)$ are normally distributed with zero mean, the second expectation in (1.32) equals zero, $\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0$, for any index i, and thus $\mathbb{E}[I(T)] = 0$.

This will be different in the case of Stratonovich' calculus, where a term $\mathbb{E}\left[g(t_{i+\frac{1}{2}})\left(W(t_{i+1}) - W(t_i)\right)\right]$ would not be equal to 0.

The Itô isometry property states that for any stochastic process $g(t)$, satisfying the usual regularity conditions, the following equality holds,

$$\boxed{\mathbb{E}\left[\left(\int_0^T g(t)\mathrm{d}W(t)\right)^2\right] = \int_0^T \mathbb{E}[g^2(t)]\mathrm{d}t.} \tag{1.33}$$

To prove this equality, we make again use of an equally spaced partitioning, $0 = t_0 < t_1 < \cdots < T = t_m$, and write

$$\begin{aligned}\int_0^T g(t)\mathrm{d}W(t) &= \lim_{m\to\infty}\sum_{i=0}^{m-1} g(t_i)\left(W(t_{i+1}) - W(t_i)\right)\\ &= \lim_{m\to\infty}\sum_{i=0}^{m-1} g(t_i)\Delta W_i,\end{aligned}\tag{1.34}$$

with $\Delta W_i := W(t_{i+1}) - W(t_i)$. The square of the integral reads:

$$\begin{aligned}\left[\int_0^T g(t)\mathrm{d}W(t)\right]^2 &= \lim_{m\to\infty}\left[\sum_{i=0}^{m-1} g(t_i)\Delta W_i\right]^2\\ &= \lim_{m\to\infty}\left[\sum_{i=0}^{m-1} g^2(t_i)\Delta W_i^2 + 2\sum_{i=0}^{m-1}\sum_{j=i+1}^{m-1} g(t_i)g(t_j)\Delta W_i\Delta W_j\right].\end{aligned}$$

We take the expectations at both sides of the equality above, and, since for $i \neq j$ ΔW_i is independent of ΔW_j (independent increments), the expectation of the double sum is equal to zero, i.e.

$$\mathbb{E}\left[\int_0^T g(t)\mathrm{d}W(t)\right]^2 = \lim_{m\to\infty}\sum_{i=0}^{m-1}\mathbb{E}\left[g^2(t_i)\Delta W_i^2\right].$$

We now use the tower property of expectations, i.e.,

$$\mathbb{E}[X|\mathcal{F}(s)] = \mathbb{E}\left[\mathbb{E}[X|\mathcal{F}(t)]|\mathcal{F}(s)\right], \quad s < t.$$

By setting $s = t_0 \equiv 0$ and $t = t_i$ in the expression above, we have:

$$\begin{aligned}\mathbb{E}\left[\int_0^T g(t)\mathrm{d}W(t)\right]^2 &= \lim_{m\to\infty}\sum_{i=0}^{m-1}\mathbb{E}\left[\mathbb{E}\left[g^2(t_i)\Delta W_i^2|\mathcal{F}(t_i)\right]|\mathcal{F}(0)\right]\\ &= \lim_{m\to\infty}\sum_{i=0}^{m-1}\mathbb{E}\left[g^2(t_i)\mathbb{E}\left[\Delta W_i^2|\mathcal{F}(t_i)\right]|\mathcal{F}(0)\right].\end{aligned}\tag{1.35}$$

Since increments of Brownian motion are independent and the variance of $W(t_{i+1}) - W(t_i)$ equals $t_{i+1} - t_i$, we find

$$\mathbb{E}\left[\left(W(t_{i+1}) - W(t_i)\right)^2|\mathcal{F}(t_i)\right] = \mathbb{E}\left[\left(W(t_{i+1}) - W(t_i)\right)^2\right] = t_{i+1} - t_i, \tag{1.36}$$

so that Equation (1.35) reads:

$$\boxed{\mathbb{E}\left[\int_0^T g(t)\mathrm{d}W(t)\right]^2 = \lim_{m\to\infty}\sum_{i=0}^{m-1}\mathbb{E}\left[g^2(t_i)|\mathcal{F}(0)\right](t_{i+1}-t_i) = \int_0^T \mathbb{E}[g^2(t)]\mathrm{d}t.}$$

which defines the Itô isometry.

The following theorem, where the proof is based on the Itô isometry, confirms the existence of the Itô integral.[1]

Theorem 1.3.3 (Existence of Itô integral) *The Itô integral, as defined in (1.24), exists.*

Proof To show the existence of the integral, it is sufficient to show that

$$I_m(T) = \int_0^T g_m(t)\mathrm{d}W(t),$$

converges to some element in L^2. With the elementary process $g_m(t)$ a (random) step function on each interval of the time partition, we can write:

$$I_m(T) = \sum_{i=0}^{m-1} g_m(t_i)(W(t_{i+1}) - W(t_i)).$$

To show that the limit in Equation (1.24) exists, we consider the following limit:

$$\lim_{n,m\to\infty}(I_n(T) - I_m(T)) = 0, \text{ in } L^2. \tag{1.37}$$

If we can show that the limit (1.37) equals 0 in L^2, this implies the existence of the limit of $I_m(T)$.

For any $n > 0$ and $m > 0$ we have:

$$\begin{aligned}\mathbb{E}\left[I_n(T) - I_m(T)\right]^2 &= \mathbb{E}\left[\int_0^T (g_n(t) - g_m(t))\,\mathrm{d}W(t)\right]^2\\ &= \int_0^T \mathbb{E}\left[g_n(t) - g_m(t)\right]^2 \mathrm{d}t,\end{aligned} \tag{1.38}$$

where the second step comes from Itô's isometry. Equation (1.38) can be rewritten as:

$$\int_0^T \mathbb{E}\left[g_n(t) - g_m(t)\right]^2 \mathrm{d}t = \int_0^T \mathbb{E}\left[(g_n(t) - g(t)) + (g(t) - g_m(t))\right]^2 \mathrm{d}t. \tag{1.39}$$

[1] For uniqueness, we refer again to standard literature.

Again using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, we have:

$$\int_0^T \mathbb{E}\left[g_n(t) - g_m(t)\right]^2 \mathrm{d}t \leq 2\int_0^T \mathbb{E}\left[g_n(t) - g(t)\right]^2 \mathrm{d}t + 2\int_0^T \mathbb{E}\left[g_m(t) - g(t)\right]^2 \mathrm{d}t.$$

Using the results from Theorem 1.3.2 we find:

$$\lim_{n\to\infty} \mathbb{E}\left[g_n(t) - g(t)\right]^2 \mathrm{d}t = 0, \text{ and } \lim_{m\to\infty} \mathbb{E}\left[g_m(t) - g(t)\right]^2 \mathrm{d}t = 0,$$

and therefore:

$$0 \leq \lim_{n,m\to\infty} \mathbb{E}\left[I_n(T) - I_m(T)\right]^2 \leq 0, \tag{1.40}$$

which implies, by the squeeze theorem of sequences, that the limit of $I_m(T)$ exists. This concludes the proof. ■

Theorem 1.3.4 (Itô integral is a martingale) *For any $g(t) \in L^2$, the stochastic integral $I(T) := \int_0^T g(t)\mathrm{d}W(t)$ is a martingale with respect to the filtration $\mathcal{F}(T), T \geq 0$.*

Proof As shown in Theorem 1.3.3, $I(T)$ exists, so $\mathbb{E}[I(T)] < \infty$. For any random process $g(t)$ in L^2, we have:

$$\begin{aligned}\mathbb{E}[I(t+\Delta t)|\mathcal{F}(t)] &= \mathbb{E}[I(t+\Delta t) - I(t) + I(t)|\mathcal{F}(t)]\\ &= \mathbb{E}[I(t+\Delta t) - I(t)|\mathcal{F}(t)] + \mathbb{E}[I(t)|\mathcal{F}(t)] = 0 + I(t),\end{aligned}$$

using the property of independent increments and the property of measurability. ■

1.3.3 Martingale representation theorem

Theorem 1.3.5 (Martingale Representation Theorem) *With $t_0 \leq t \leq T$, let $W(t)$, with $W(t_0) = W_0$, be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{F}(t)$, a filtration generated by the Brownian motion. With $X(t)$, a martingale relative to this filtration, there is an adapted process $g(t)$, such that:*

$$\mathrm{d}X(t) = g(t)\mathrm{d}W(t), \quad \textit{or} \quad X(t) = X_0 + \int_0^t g(z)\mathrm{d}W(z).$$

For a proof, see Øksendal [2000].

The theorem above states that if the process $X(t)$ is a martingale, adapted to the filtration generated by the Brownian motion $W^{\mathbb{P}}(t)$, then process $X(t)$ needs to be of the following form

$$\boxed{\mathrm{d}X(t) = g(t)\mathrm{d}W^{\mathbb{P}}(t),}$$

for some process $g(t)$.

As the integral formulation, in (1.23), is equivalent to the SDE in (1.22), we can conclude that an SDE without any drift term is a martingale.

Example 1.3.1 (Solution of Itô integral) In this example, we solve the following stochastic integral,

$$I(T) = \int_0^T W(t)\mathrm{d}W(t). \tag{1.41}$$

Following Definition 1.3.1, we have $g(t) := W(t)$, consider an equally spaced partition $0 = t_0 < t_1 < \cdots < t_m = T$, with $t_i = i\frac{T}{m}$, with an equidistant time increment $\Delta t = (t_{i+1} - t_i)$ and define a sequence of some elementary functions satisfying (1.25), as follows:

$$g_m(t) = \begin{cases} W(0), & \text{for} \quad 0 \le t < t_1, \\ W(t_1), & \text{for} \quad t_1 \le t < t_2, \\ \dots & \\ W(t_m), & \text{for} \quad t_{m-1} \le t < t_m. \end{cases} \tag{1.42}$$

Let us verify for this sequence of elementary functions that the condition in (1.25) is satisfied, i.e.

$$\begin{aligned} \lim_{m\to\infty} \mathbb{E}\left[\int_0^T (g_m(t) - g(t))^2 \,\mathrm{d}t\right] &= \lim_{m\to\infty} \mathbb{E}\left[\int_0^T (g_m(t) - W(t))^2 \,\mathrm{d}t\right] \\ &= \lim_{m\to\infty} \sum_{i=0}^{m-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[W(t_i) - W(t)\right]^2 \mathrm{d}t. \end{aligned}$$

As $t_i < t$, the last term can be written as

$$\begin{aligned} \lim_{m\to\infty} \sum_{i=0}^{m-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[W(t_i) - W(t)\right]^2 \mathrm{d}t &= \lim_{m\to\infty} \sum_{i=0}^{m-1} \int_{t_i}^{t_{i+1}} (t - t_i)\mathrm{d}t \\ &= \lim_{m\to\infty} \sum_{i=0}^{m-1} \frac{1}{2}\left(t_{i+1} - t_i\right)^2. \end{aligned}$$

Using the fact that $\Delta t = t_{i+1} - t_i$, we have:

$$\lim_{m\to\infty} \sum_{i=0}^{m-1} \frac{1}{2}(\Delta t)^2 = 0. \tag{1.43}$$

This is because

$$\lim_{m\to\infty}\sum_{i=0}^{m-1}\frac{1}{2}(\Delta t)^2=\lim_{m\to\infty}\sum_{i=0}^{m-1}\frac{1}{2}\left[(i+1)\frac{T}{m}-i\frac{T}{m}\right]^2=\lim_{m\to\infty}\frac{1}{2}\sum_{i=0}^{m-1}\frac{T^2}{m^2},$$

which converges to 0 for $m\to\infty$.

So, the condition in (1.25) holds, and we continue with the discrete version of the integral in (1.41):

$$\int_0^T W(t)\mathrm{d}W(t)=\lim_{m\to\infty}\sum_{i=0}^{m-1}W(t_i)\left(W(t_{i+1})-W(t_i)\right). \tag{1.44}$$

By basic algebra[2] the right-hand side of (1.44) can be simplified to:

$$\begin{aligned}\sum_{i=0}^{m-1}W(t_i)\left(W(t_{i+1})-W(t_i)\right)&=\frac{1}{2}\sum_{i=0}^{m-1}\left(W^2(t_{i+1})-W^2(t_i)\right)\\&\quad-\frac{1}{2}\sum_{i=0}^{m-1}\left(W(t_{i+1})-W(t_i)\right)^2.\end{aligned}$$

The first element at the right-hand side of the expression above can be recognized as a *telescopic sum*, i.e.

$$\begin{aligned}\sum_{i=0}^{m-1}\left(W^2(t_{i+1})-W^2(t_i)\right)&=\cancel{W^2(t_1)}-W^2(t_0)+\cancel{W^2(t_2)}-\cancel{W^2(t_1)}+\cdots\\&=W^2(t_m)-W^2(t_0),\end{aligned}$$

as $t_m=T$, $t_0=0$ and $W^2(t_0)\equiv 0$, giving the following simplification:

$$\lim_{m\to\infty}\sum_{i=0}^{m-1}W(t_i)(W(t_{i+1})-W(t_i))=\frac{1}{2}W^2(T)-\frac{1}{2}\lim_{m\to\infty}\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2. \tag{1.45}$$

To finalize the task of calculating the integral in (1.41), we calculate the sum at the right-hand side of (1.45) for which we take the expectation:

$$\begin{aligned}\lim_{m\to\infty}\mathbb{E}\left[\sum_{i=0}^{m-1}\left(W(t_{i+1})-W(t_i)\right)^2\right]&=\lim_{m\to\infty}\sum_{i=0}^{m-1}\mathbb{E}\left[W(t_{i+1})-W(t_i)\right]^2\\&=\lim_{m\to\infty}\sum_{i=0}^{m-1}(t_{i+1}-t_i)=T.\end{aligned} \tag{1.46}$$

[2] $x(y-x)=\frac{1}{2}(y^2-x^2)-\frac{1}{2}(y-x)^2$.

We need to verify whether T is the limit of the summation in L^2. This, by definition, can be confirmed by the following expectation:

$$\begin{aligned}
&\lim_{m\to\infty} \mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2 - T\right]^2 \\
&\quad= \lim_{m\to\infty} \mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2\right]^2 \\
&\qquad - 2T\lim_{m\to\infty} \mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2\right] + T^2.
\end{aligned}$$

Using the results from (1.46), we find:

$$\lim_{m\to\infty} \mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2 - T\right]^2 = \lim_{m\to\infty} \mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2\right]^2 - T^2.$$

The expectation at the right-hand side can be further simplified, as follows,

$$\mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2\right]^2 = \mathbb{E}\left[\sum_{i=0}^{m-1}(\Delta t)^2 Z_i^4 + 2\sum_{i=0}^{m-1}\sum_{j=i+1}^{m-1}(\Delta t)^2 Z_i^2 Z_j^2\right].$$

with $Z_i = \mathcal{N}(0,1)$. Because the fourth moment of a standard normal random variable equals 3, i.e. $\mathbb{E}[Z_i^4] = 3$, and any Z_i is mutually independent of Z_j, for $i \neq j$, we get

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=0}^{m-1}(\Delta t)^2 Z_i^4 + 2\sum_{i=0}^{m-1}\sum_{j=i+1}^{m-1}(\Delta t)^2 Z_i^2 Z_j^2\right] &= 3\sum_{i=0}^{m-1}(\Delta t)^2 + 2\sum_{i=0}^{m-1}\sum_{j=i+1}^{m-1}(\Delta t)^2 \\
&= 3(\Delta t)^2 m + (\Delta t)^2(m^2 - m) \\
&= T^2 + 2T\frac{1}{m}.
\end{aligned}$$

In the limit case $m \to \infty$, we find:

$$\lim_{m\to\infty} \mathbb{E}\left[\sum_{i=0}^{m-1}(W(t_{i+1})-W(t_i))^2 - T\right]^2 = \lim_{m\to\infty}\left[T^2 + 2T\frac{1}{m} - T^2\right] = 0,$$

implying convergence of the sum in (1.46) to T, in L^2.

This result implies the following solution for the summation in (1.45):

$$\lim_{m\to\infty}\sum_{i=0}^{m-1} W(t_i)\left(W(t_{i+1})-W(t_i)\right) = \frac{1}{2}W^2(T) - \frac{1}{2}T. \tag{1.47}$$

and the solution of the integral in (1.41) is therefore given by:

$$\boxed{\int_0^T W(t)\mathrm{d}W(t) = \frac{1}{2}W^2(T) - \frac{1}{2}T.} \tag{1.48}$$

♦

We conclude this introductory part with the following short summary.

The Itô integral defined by (1.23) has a number of important properties, like

a. For every time $t \geq 0$, $I(t)$ is $\mathcal{F}(t)$-measurable,

b. $\mathbb{E}[I(t)|\mathcal{F}(0)] = 0$,

c. $\mathbb{E}[I(t)|\mathcal{F}(s)] = I(s)$, for $s < t$, which is the martingale property, and the dynamics $\mathrm{d}I(t)$ do not contain a drift term,

d. Itô isometry:

$$\mathbb{E}\left[\int_0^T g(t)\mathrm{d}W(t)\right]^2 = \int_0^T \mathbb{E}[g^2(t)]\mathrm{d}t,$$

e. For $0 \leq a < b < c$, it follows that

$$\int_a^c g(t)\mathrm{d}W(t) = \int_a^b g(t)\mathrm{d}W(t) + \int_b^c g(t)\mathrm{d}W(t),$$

f. Another equality which holds true is:

$$\int_a^c (\alpha \cdot g(t) + h(t))\mathrm{d}W(t) = \alpha \int_a^c g(t)\mathrm{d}W(t) + \int_a^c h(t)\mathrm{d}W(t),$$

with $a < c$, $\alpha \in \mathbb{R}$, and $h(t) \in \mathcal{F}(t)$.

Proofs for these statements are given in stochastic calculus textbooks, such as [Shreve, 2004].

1.4 Exercise set

Exercise 1.1 With

$$F_{\mathcal{N}(0,1)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}}\, dz,$$

show that

$$F_{\mathcal{N}(0,1)}(x) + F_{\mathcal{N}(0,1)}(-x) = 1.$$

Exercise 1.2 Use $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$, and $\mathbb{E}[\alpha X] = \alpha\mathbb{E}[X] \quad (\alpha \in \mathbb{R})$, to show,

$$\mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2;$$

and show with this, that,

$$\mathbb{V}\text{ar}[\alpha X] = \alpha^2 \mathbb{V}\text{ar}[X], \text{ with } \alpha \in \mathbb{R}.$$

Exercise 1.3 Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = a + bX$ $(b \neq 0)$. Determine $\mathbb{E}[Y]$, $\mathbb{V}\text{ar}[Y]$ and the distribution function of Y. Determine $\mathbb{E}[e^X]$.

Exercise 1.4 Show that the standard normal distribution function, $F_{\mathcal{N}(0,1)}(x)$, can be calculated with the so-called *error function*,

$$\text{erf }(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-s^2}\, ds,$$

with the help of the formula,

$$F_{\mathcal{N}(0,1)}(x) = \frac{1 + \text{ erf}\left(\frac{x}{\sqrt{2}}\right)}{2}.$$

Exercise 1.5 With $X_1, \ldots, X_n$ i.i.d. random variables with the same distribution, expectation μ and variance σ^2. Random variable $\bar{X}$ is defined as the arithmetic average of these variables, which is the *sample mean*,

$$\bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k.$$

a. Show that $\mathbb{E}[\bar{X}] = \mu$.
b. Show that $\mathbb{V}\text{ar}[\bar{X}] = \sigma^2/n$. The random variable $\bar{\upsilon}_N^2$, which is defined as:

$$\bar{\upsilon}_N^2 := \frac{\sum_{k=1}^{N} (X_k - \bar{X})^2}{N-1},$$

is the *sample variance*.
c. Show that $\sum_{k=1}^{N} (X_k - \bar{X})^2 = \sum_{k=1}^{N} X_k^2 - N\bar{X}^2$.
d. Show that $\mathbb{E}[\bar{\upsilon}_N^2] = \sigma^2$.

Exercise 1.6 For a given Brownian motion $W(t)$:

a. Solve analytically the expectation,

$$\mathbb{E}\left[W^4(t) - \frac{1}{2}W^3(t)\right].$$

b. Find analytically $\mathbb{V}\mathrm{ar}[Z(t)]$, with

$$Z(t) = W(t) - \frac{t}{T}W(T-t), \quad \text{for } 0 \leq t \leq T.$$

Exercise 1.7 Show theoretically that

$$\int_0^t W(z)\mathrm{d}z = \int_0^t (t-z)\mathrm{d}W(z).$$

Exercise 1.8 With standard Brownian motion $W(t),\ t \geq 0$, calculate the integral

$$\int_{z=0}^{T}\int_{s=0}^{z} \mathrm{d}W(s)\mathrm{d}W(z).$$

Exercise 1.9 Show that, for a continuously, differentiable function $g(t)$, the process

$$X(t) = g(t)W(t) - \int_0^t \frac{\mathrm{d}g(z)}{\mathrm{d}z}W(z)\mathrm{d}z,$$

is a martingale, and subsequently show that

$$\mathbb{E}[\mathrm{e}^{2t}W(t)] = \mathbb{E}\left[\int_0^t 2\mathrm{e}^{2z}W(z)\mathrm{d}z\right].$$

Exercise 1.10 A time continuous stochastic process $\{X(t);\ t \in \mathcal{T}\}$ is called a Gaussian process, if for any set of time indices, $t_1, \ldots, t_m$, all linear combinations of $(X(t_1), \ldots, X(t_m))$ are governed by a univariate normal distribution.

Given a Gaussian process $X(t),\ t > 0$ with $X(0) = 0$. Determine the following covariance,

$$2\mathbb{C}\mathrm{ov}[X(s), X(t)] = \mathbb{E}[X^2(s)] + \mathbb{E}[X^2(t)] - \mathbb{E}[(X(t) - X(s))^2],\ 0 < s < t.$$

Exercise 1.11 Consider the following SDE,

$$\mathrm{d}X(t) = \mu\mathrm{d}t + \sigma\mathrm{d}W(t), \quad X(t_0) = x_0, \tag{1.49}$$

with some constants μ and σ. Show that, by choosing $t_0 = 0$, the integrated process $X(t)$ follows the following distribution:

$$\int_0^T X(t)\mathrm{d}t \sim \mathcal{N}\left(x_0T + \frac{1}{2}\mu T^2, \frac{1}{3}\sigma^2T^3\right).$$

CHAPTER 2

Introduction to Financial Asset Dynamics

In this chapter:

In **Section 2.1**, we present the mathematical basis of stochastic models for financial asset prices. In particular, we focus on the *Geometric Brownian Motion stock price model. Itô's lemma* is discussed in this chapter, as it plays an important role for many derivations.
In **Section 2.2** some first variations to the basic geometric Brownian motion process are presented. The *martingale property of financial asset prices* is explained in **Section 2.3**.

Keywords: model for asset prices, geometric Brownian motion, martingales.

2.1 Geometric Brownian motion asset price process

A stock or share is a financial asset, which represents ownership of a tiny piece of a company, and is traded on financial markets. According to the efficient market hypothesis, the stock price is determined by the present value of the company plus the expectations of the companies' future performance. These expectations give rise to uncertainty in the asset price, as seen in the form of bid and ask prices offered by the participants in the financial markets. Asset prices thus have an element of randomness, commonly modeled by *stochastic differential equations* (SDEs). Closed-form solutions to these SDEs are available only in exceptional cases. They can serve as a validation for numerical techniques, or as building blocks for SDE asset price models of increasing complexity.

The most commonly used asset price process in finance is the geometric Brownian Motion (GBM) model, where the *logarithm of the asset price* follows an arithmetic Brownian motion, driven by a Wiener process $W(t)$.

The asset price $S(t)$ is said to follow a GBM process, when it satisfies the following SDE:

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{P}}(t), \quad \text{with} \quad S(t_0) = S_0, \tag{2.1}$$

where the Brownian motion $W^{\mathbb{P}}(t)$ is under the so-called real-world measure $\mathbb{P}$, $\mu = \mu^{\mathbb{P}}$ denotes the *drift parameter*, i.e. a constant deterministic growth rate of the stock, and σ is the (constant) percentage volatility parameter. Model (2.1) is also referred to as the *Samuelson model*. This is a short-hand notation for the integral formulation,

$$S(t) = S_0 + \int_{t_0}^{t} \mu S(z)\mathrm{d}z + \int_{t_0}^{t} \sigma S(z)\mathrm{d}W^{\mathbb{P}}(z). \tag{2.2}$$

The amount by which an asset price differs from its expected value is determined by the volatility parameter σ. Volatility is thus a statistical measure of the tendency of an asset to rise or fall sharply within a period of time. It can be calculated, for example, by the variance of the asset prices, measured within a certain time period. A high volatile market implies that prices have large deviations from their mean value in short periods of time. Although both μ and σ in (2.1) are assumed to be constant, the extension to time-dependent functions is also possible. It should however be noticed that these parameter values are estimates for the growth rate and the volatility *in the future*, i.e., for $t > t_0$.

Before we dive deeper into our first financial asset stochastic process, we have a look at the (deterministic) money savings account, which we denote by $M(t)$.

Definition 2.1.1 (Markov process) *A stochastic process is a Markov process, if the conditional probability distribution of future states depends only on the present state, and not on the history. In a financial setting, this implies that we assume that the current stock price contains all information of the past asset prices. The adapted stock price process $S(t)$ on a filtered probability space has the Markov property, if for each bounded and measurable function $g : \mathbb{R}^N \to \mathbb{R}$,*

$$\mathbb{E}[g(S(t))|\mathcal{F}(s)] = \mathbb{E}[g(S(t))|S(s)], \quad s \leq t. \tag{2.3}$$

◀

Definition 2.1.2 (Money-savings account) *The simplest concept in finance is the time value of money. One unit of currency today is worth more than one unit in a year's time, when interest rates are positive, and less than one unit in the case of negative interest rates.. Particularly, we will focus our attention on* compounded interest, *which is defined as interest on earlier interest payments (on an initial notional amount). This interest can be either discretely or continuously compounded. Receiving m discrete interest payments at a rate of r/m per year, on an initial notional $M(0) = 1$, gives, after one year $T = 1$, the* *amount,*

$$M(T) = \left(1 + \frac{r}{m}\right)^m.$$

When the interest payments come in increasingly smaller time intervals, at a proportionally smaller interest rate (take the limit $m \to \infty$), this defines the continuous *interest rate. It can be shown, by basic calculus, that*

$$\lim_{m\to\infty}\left(1 + \frac{r}{m}\right)^m = \mathrm{e}^{m \log\left(1+\frac{r}{m}\right)} = \mathrm{e}^r.$$

At a time point t, we will have the amount e^{rt} on the bank account. This is the so-called money-savings account.

With an amount $M(t)$ at time t on the money-savings account, the Taylor expansion indicates that at time point $M(t + \Delta t)$, the amount will have increased by,

$$M(t + \Delta t) - M(t) \approx \frac{\mathrm{d}M}{\mathrm{d}t}\Delta t + \ldots.$$

The change in money is proportional to the initial amount, the interest rate r and the time period the money is on the account, giving,

$$\frac{\mathrm{d}M(t)}{\mathrm{d}t} = rM(t) \ . \tag{2.4}$$

Starting with $M(0) = 1$ at $t_0 = 0$, we will have $M(t) = \mathrm{e}^{rt}$, at time point t. On the other hand, if we wish to receive $M(T) = 1$ at a future time point $t = T$, we need to put the amount $\mathrm{e}^{-r(T-t)}$ on our money-savings account at time point $t_0 = 0$. Clearly, money grows exponentially *in continuous time, however, with a very small (positive) interest rate r, the growth is still rather limited in the long run.*

Compared to the stochastic GBM asset dynamics in Equation (2.1), we find a drift term in the deterministic interest rate dynamics in (2.4) with $\mu = r$, and $\sigma = 0$. Stochastic interest rates, in the form of so-called short-rates and Libor rates, will also be discussed in detail in this book, starting from Chapter 11. These are important stochastic models for example for options on interest rates. ◀

2.1.1 Itô process

Itô's lemma is fundamental for stochastic processes, as it enables us to handle the Wiener increment $\mathrm{d}W(t)$ as in (2.1), when $\mathrm{d}t \to 0$ (similar to a Taylor expansion for deterministic variables and functions). By Itô's lemma we can derive solutions to SDEs, and we can derive pricing partial differential equations (PDEs) for financial derivative products in the subsequent chapter. We first discuss some issues related to the so-called *Itô processes.*

Definition 2.1.3 (Itô process) *Let us consider the following SDE, corresponding to the Itô process $X(t)$,*

$$\mathrm{d}X(t) = \bar{\mu}(t, X(t))\mathrm{d}t + \bar{\sigma}(t, X(t))\mathrm{d}W(t), \quad \textit{with} \quad X(t_0) = X_0, \tag{2.5}$$

with two general functions for the drift $\bar{\mu}(t,x)$ and the volatility $\bar{\sigma}(t,x)$. These two functions cannot be "just any" functions; they need to satisfy the following two Lipschitz conditions:

$$|\bar{\mu}(t,x) - \bar{\mu}(t,y)|^2 + |\bar{\sigma}(t,x) - \bar{\sigma}(t,y)|^2 \leq K_1|x-y|^2,$$
$$|\bar{\mu}(t,x)|^2 + |\bar{\sigma}(t,x)|^2 \leq K_2(1+|x|^2),$$

for some constants $K_1, K_2 \in \mathbb{R}^+$ and x and y in $\mathbb{R}$. The two conditions above state that the drift and volatility terms should not increase too rapidly. When these conditions hold, then, with probability one, a continuous, adapted solution of (2.5) exists, and the solution satisfies $\sup_{0\leq t\leq T}\mathbb{E}[X^2(t)] < \infty$. ◀

2.1.2 Itô's lemma

With stochastic process $X(t)$ determined by (2.5), another process $Y(t)$ can be defined as a function of t and $X(t)$, i.e., $Y(t) := g(t,X)$.[1] $Y(t)$ is a stochastic process and its SDE can also be determined. The procedure to derive the SDE for process $Y(t)$ is given by Itô's lemma below.

Before we present Itô's lemma however, we give some heuristics. To derive the dynamics $\mathrm{d}Y(t)$ for $Y(t) = g(t,X)$, we may take a look at the 2D Taylor series expansion around some point (t_0, X_0), i.e.,

$$\begin{aligned} g(t,X) = g(t_0,X_0) &+ \left.\frac{\partial g(t,X)}{\partial t}\right|_{t=t_0}\Delta t + \frac{1}{2}\left.\frac{\partial^2 g(t,X)}{\partial t^2}\right|_{t=t_0}(\Delta t)^2 \\ &+ \left.\frac{\partial g(t,X)}{\partial X}\right|_{X=X_0}\Delta X + \frac{1}{2}\left.\frac{\partial^2 g(t,X)}{\partial X^2}\right|_{X=X_0}(\Delta X)^2 \\ &+ \left.\frac{\partial^2 g(t,X)}{\partial t \partial X}\right|_{X=X_0, t=t_0}\Delta X \Delta t + \cdots, \end{aligned} \tag{2.6}$$

with $\Delta t = t - t_0$, and $\Delta X = X - X_0$. For $t \to t_0$ and $X \to X_0$, and $\mathrm{d}t = \lim_{t\to t_0} t - t_0$, $\mathrm{d}X = \lim_{X\to X_0} X - X_0$, we may write (2.6) as follows:

$$\begin{aligned} \mathrm{d}g(t,X) = \frac{\partial g}{\partial t}\mathrm{d}t &+ \frac{1}{2}\frac{\partial^2 g}{\partial t^2}(\mathrm{d}t)^2 + \frac{\partial g}{\partial X}\mathrm{d}X + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}\mathrm{d}X^2 \\ &+ \frac{\partial^2 g}{\partial t \partial X}\mathrm{d}X\mathrm{d}t + \cdots. \end{aligned} \tag{2.7}$$

[1] Here $X = X(t)$ serves as an independent variable.

In (2.7) we encounter infinitely many terms. Many of those terms however can be neglected in the limit $\mathrm{d}t \to 0$. When the time increment $\mathrm{d}t$ goes to 0, terms $(\mathrm{d}t)^2$ tend to 0 much faster than the terms with $\mathrm{d}t$. This holds true for any term $(\mathrm{d}t)^n$, with $n > 1$. Conventionally, this convergence behavior is described by the little-o notation, i.e. $(\mathrm{d}t)^2 = o(\mathrm{d}t)$.

It is common to denote $(\mathrm{d}t)^2 = 0$, but we need to keep in mind that this equality actually means "order $\mathrm{d}t \to 0$".

Remark 2.1.1 (Little-*o* and big-*O* notation) *By definition,*

$$g(x) = O(h(x)), \quad \textit{if} \;\; |g(x)| < c \cdot h(x),$$

for some constant "c" and for sufficiently large x. Little-o describes the following asymptotic limit,

$$g(x) = o(h(x)), \quad \textit{if} \;\; \lim_{x \to \infty} \frac{g(x)}{h(x)} = 0.$$

Informally, big-O can be thought of as "$g(x)$ does not grow faster than $h(x)$" and of little-o as "$g(x)$ grows much slower than $h(x)$". ▲

Example 2.1.1 (Little-*o* and big-*O*) As an example, the function $g(x) = 5x^2 - 1x + 9$ is $O(x^2)$, but it is not $o(x^2)$, because $\lim_{x \to \infty} \frac{g(x)}{x^2} = 5$.

The following statements are true for little-o,

$$x^2 = o(x^3), \;\; x^2 = o(x!), \;\; \log(x^2) = o(x).$$

The following statements are true under big-O notation,

$$x^2 = O(x^2), \;\; x^2 = O(x^2 + x), \;\; x^2 = O(100x^2),$$

however, they are not true when little-o would be used. ♦

Itô's table

The equality in (2.7) can be simplified, by neglecting the higher-order $\mathrm{d}t$-terms, by writing,

$$\mathrm{d}g(t, X) = \frac{\partial g}{\partial t}\mathrm{d}t + \frac{\partial g}{\partial X}\mathrm{d}X + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}(\mathrm{d}X)^2. \tag{2.8}$$

We need to make statements about the term $\mathrm{d}X\mathrm{d}X$, which, in the case of Equation (2.5), reads:

$$(\mathrm{d}X)^2 = \bar{\mu}^2(t, X)(\mathrm{d}t)^2 + \bar{\sigma}^2(t, X)(\mathrm{d}W)^2 + 2\bar{\mu}(t, X)\bar{\sigma}(t, X)\mathrm{d}W\mathrm{d}t. \tag{2.9}$$

Two specific terms $\mathrm{d}t\mathrm{d}W$ and $\mathrm{d}W\mathrm{d}W$ need to be determined. The expectation of $\mathrm{d}t\mathrm{d}W$ equals 0 (because an expectation of a Brownian increment, scaled by a constant, equals zero) and the standard deviation equals $\mathrm{d}t^{\frac{3}{2}}$ (as the standard deviation of $\mathrm{d}W$ is equal to $\sqrt{\mathrm{d}t}$). We have $(\mathrm{d}t)^{\frac{3}{2}}$, implying that $\mathrm{d}t\mathrm{d}W$ goes to

Table 2.1: *Itô multiplication table for Wiener process.*

	$\mathrm{d}t$	$\mathrm{d}W(t)$
$\mathrm{d}t$	0	0
$\mathrm{d}W(t)$	0	$\mathrm{d}t$

0 rapidly when $\mathrm{d}t \to 0$. Regarding the other term, the expectation of $\mathrm{d}W\mathrm{d}W$ is equal to $\mathrm{d}t$ implying that $\mathrm{d}W\mathrm{d}W$ is of order $\mathrm{d}t$ when $\mathrm{d}t \to 0$.

So, when deriving the Itô dynamics, we make use of the Itô multiplication table, where the cross terms involving the Wiener process are handled as in Table 2.1, see also the discussion in Privault [1998].

Remark 2.1.2 *The approximation* $(\mathrm{d}W)^2 = \mathrm{d}t$.
Let us consider the following expectation, $\mathbb{E}\left[(\mathrm{d}W)^2\right]$. *By the same steps as in the derivations of the Itô isometry, see* (1.36), *we have*

$$\mathbb{E}\left[(\mathrm{d}W)^2\right] = \lim_{\Delta t \to 0} \mathbb{E}\left[(W(t+\Delta t) - W(t))^2\right] = \lim_{\Delta t \to 0} \Delta t = \mathrm{d}t, \quad (2.10)$$

and the variance is equal to:

$$\begin{aligned}
\mathbb{V}ar\left[(\mathrm{d}W)^2\right] &= \lim_{\Delta t \to 0} \mathbb{V}ar\left[(W(t+\Delta t) - W(t))^2\right] \\
&= \lim_{\Delta t \to 0} \mathbb{E}\left[(W(t+\Delta t) - W(t))^4\right] \\
&- \lim_{\Delta t \to 0} \left(\mathbb{E}\left[(W(t+\Delta t) - W(t))^2\right]\right)^2 \\
&= \lim_{\Delta t \to 0} 3(\Delta t)^2 - \lim_{\Delta t \to 0} (\Delta t)^2 = \lim_{\Delta t \to 0} 2(\Delta t)^2 \\
&= 2(\mathrm{d}t)^2.
\end{aligned}$$

We conclude that the variance of $(\mathrm{d}W)^2$ *converges to zero much faster than the expectation, when* $\Delta t \to 0$. *Because of this, we have as a stochastic calculus rule,*

$$(\mathrm{d}W)^2 = \mathrm{d}t,$$

as the variance approaches zero rapidly in the limit. ▲

We can therefore write (2.9) as:

$$\boxed{(\mathrm{d}X)^2 \approx \bar{\sigma}^2(t, X)\mathrm{d}t.}$$

By collecting all building blocks we will write the dynamics of $g(t, X)$ as follows:

$$\mathrm{d}g(t, X) = \frac{\partial g}{\partial t}\mathrm{d}t + \left(\bar{\mu}(t, X)\frac{\partial g}{\partial X} + \frac{1}{2}\bar{\sigma}^2(t, X)\frac{\partial^2 g}{\partial X^2}\right)\mathrm{d}t + \frac{\partial g}{\partial X}\bar{\sigma}(t, X)\mathrm{d}W(t).$$

We now have the following lemma:

Theorem 2.1.1 (Itô's lemma) *Suppose a process $X(t)$ follows the Itô dynamics,*

$$\mathrm{d}X(t) = \bar{\mu}(t, X(t))\mathrm{d}t + \bar{\sigma}(t, X(t))\mathrm{d}W(t), \text{ with } X(t_0) = X_0,$$

where drift $\bar{\mu}(t, X(t))$ and diffusion $\bar{\sigma}(t, X(t))$ satisfy the standard Lipschitz conditions on the growth of these functions (as in Definition 2.1.3*).*
Let $g(t, X)$ be a function of $X = X(t)$ and time t, with continuous partial derivatives, $\partial g/\partial X$, $\partial g^2/\partial X^2$, $\partial g/\partial t$. A stochastic variable $Y(t) := g(t, X)$ then also follows an Itô process, governed by the same Wiener process $W(t)$, i.e.,

$$\mathrm{d}Y(t) = \left(\frac{\partial g}{\partial t} + \bar{\mu}(t, X)\frac{\partial g}{\partial X} + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}\bar{\sigma}^2(t, X)\right)\mathrm{d}t + \frac{\partial g}{\partial X}\bar{\sigma}(t, X)\mathrm{d}W(t).$$

The formal proof of this theorem is somewhat involved, and can be found in several textbooks on stochastic processes [Shreve, 2004].

Again, Itô's lemma above is a short-hand notation for the integral formulation,

$$\begin{aligned} Y(t) = Y_0 + \int_{t_0}^{t} \mu S(z)\left(\frac{\partial g}{\partial z} + \bar{\mu}(z, X)\frac{\partial g}{\partial X} + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}\bar{\sigma}^2(z, X)\right)\mathrm{d}z \\ + \int_{t_0}^{t} \frac{\partial g}{\partial X}\bar{\sigma}(z, X)\mathrm{d}W(z). \end{aligned} \tag{2.11}$$

Itô's lemma can be used to find the solutions of a number of interesting integrals and SDEs.

Example 2.1.2 (Solution of Itô integral by Itô calculus) Recall the stochastic integral in (1.41) in Example 1.3.1, $\int_0^T W(t)\mathrm{d}W(t)$. This integral can also be solved with the help of Itô's calculus. Consider the basic stochastic process, i.e., $X(t) = W(t)$, which does not have a drift term and its volatility coefficient equals 1, so $\mathrm{d}X(t) = 0 \cdot \mathrm{d}t + 1 \cdot \mathrm{d}W(t)$.

If we apply Itô's lemma to $g(X(t)) = X^2(t)$, we find:

$$\begin{aligned} \mathrm{d}g(X) &= \frac{\partial g}{\partial X}\mathrm{d}X(t) + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}(\mathrm{d}X(t))^2 \\ &= 2X(t)\mathrm{d}X(t) + (\mathrm{d}X(t))^2. \end{aligned}$$

After substitution of $W(t) = X(t)$, we find for the dynamics of $g(X(t)) = W^2(t)$,

$$\mathrm{d}W^2(t) = 2W(t)\mathrm{d}W(t) + (\mathrm{d}W(t))^2.$$

With $\mathrm{d}W(t)\mathrm{d}W(t) = \mathrm{d}t$ and integration of both sides, we arrive at:

$$\int_0^T \mathrm{d}W^2(t) = 2\int_0^T W(t)\mathrm{d}W(t) + \int_0^T \mathrm{d}t,$$

which is equivalent to:

$$\int_0^T W(t)\mathrm{d}W(t) = \frac{1}{2}\int_0^T \mathrm{d}W^2(t) - \frac{1}{2}\int_0^T \mathrm{d}t = \frac{1}{2}W^2(T) - \frac{1}{2}T,$$

providing indeed the solution to the integral of interest.

Although the technique presented above is elegant and shorter than dealing with the partitioned domains, as in (1.42), it is not always easy to find a function $g(\cdot)$ so that Itô's lemma can be applied. ♦

Example 2.1.3 (Stochastic integral $\int_0^T W(t)\mathrm{d}t$) We show here that

$$I(T) = \int_0^T W(t)\mathrm{d}t = \int_0^T (T-t)\mathrm{d}W(t).$$

By using $X(t) = W(t)$ and applying Itô's lemma to $g(t, X(t)) = (T-t)X(t)$, we find:

$$\begin{aligned}\mathrm{d}g(X) &= \frac{\partial g}{\partial t}\mathrm{d}t + \frac{\partial g}{\partial X}\mathrm{d}X(t)\\ &= -X(t)\mathrm{d}t + (T-t)\mathrm{d}X(t).\end{aligned}$$

Integrating both sides, with $X(t) = W(t)$ gives us:

$$\int_0^T \mathrm{d}((T-t)W(t)) = -\int_0^T W(t)\mathrm{d}t + \int_0^T (T-t)\mathrm{d}W(t).$$

Clearly, $\int_0^T \mathrm{d}((T-t)W(t)) = (T-T)W(T) - W(T-0)W(0) = 0$, so

$$\int_0^T W(t)\mathrm{d}t = \int_0^T (T-t)\mathrm{d}W(t). \tag{2.12}$$

♦

In Figure 2.1 we present a path of the Brownian motion $W(t)$, as well as two integral values, one from Example 2.1.2 and another from the example above.

2.1.3 Distributions of $S(t)$ and $\log S(t)$

By Itô's lemma we can show that the random variable $S := S(t)$ in (2.1) is from a *lognormal distribution*, i.e. $\log S$ is normally distributed. Using $g(t, S) = X(t) = \log S$, we obtain $\mathrm{d}g/\mathrm{d}S = 1/S$ and $\mathrm{d}^2g/\mathrm{d}S^2 = -1/S^2$, so that Itô's lemma gives us,

$$\mathrm{d}g(t,S) = \mathrm{d}X(t) = \left(\mu - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}W^{\mathbb{P}}(t), \quad g(t_0) = \log S_0. \tag{2.13}$$

As the Wiener increment $\mathrm{d}W(t)$ is normally distributed, with expectation 0 and variance $\mathrm{d}t$, Equation (2.13) confirms that $\mathrm{d}g(t, S)$ is normally distributed, with expectation $(\mu - \frac{\sigma^2}{2})\mathrm{d}t$ and variance $\sigma^2\mathrm{d}t$. The stochastic variable $Y(t) = g(t, S)$ represents a sum of the *increments* $\mathrm{d}g$ (in the limit an infinite sum can be represented by an integral), so that $Y(t) = g(t, S)$ is normally distributed, with expectation $\log S_0 + (\mu - \frac{\sigma^2}{2})(t - t_0)$ and variance $\sigma^2(t - t_0)$.

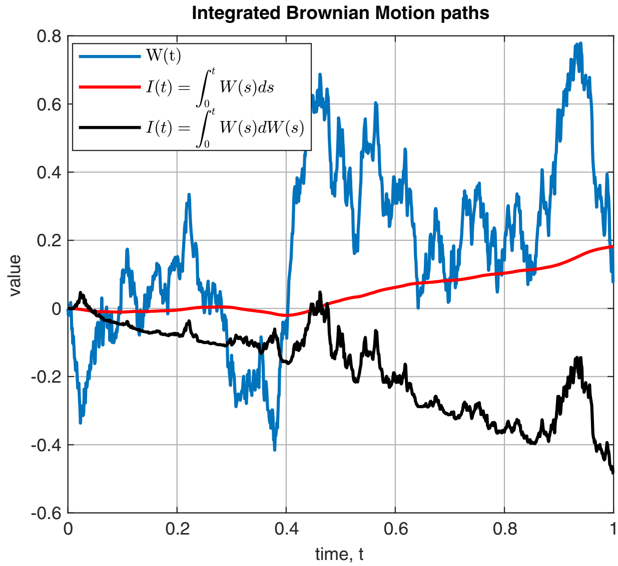


Figure 2.1: *Stochastic paths for* $W(t)$, $\int_0^t W(s)\mathrm{d}s$ *and* $\int_0^t W(s)\mathrm{d}W(s)$, *as a function of running time* t.

Example 2.1.4 (The BM and GBM distributions in time) Under the log-transformation, $X(t) = \log S(t)$, the right-hand side of the SDE (2.13) does not depend on $X(t)$, and therefore we can simply integrate both sides of the SDE,

$$\int_{t_0}^{T} \mathrm{d}X(t) = \int_{t_0}^{T} \left(\mu - \frac{1}{2}\sigma^2\right) \mathrm{d}t + \int_{t_0}^{T} \sigma \mathrm{d}W^{\mathbb{P}}(t), \tag{2.14}$$

which yields the following solution,

$$X(T) = X(t_0) + \left(\mu - \frac{1}{2}\sigma^2\right)(T - t_0) + \sigma W^{\mathbb{P}}(T - t_0). \tag{2.15}$$

Under the log-transformation, $X(T)$ is a normally distributed random variable with the following parameters:

$$X(T) \sim \mathcal{N}\left(X(t_0) + \left(\mu - \frac{1}{2}\sigma^2\right)(T - t_0), \sigma^2(T - t_0)\right).$$

With a back transformation, $S(T)$ is indeed the lognormally distributed random variable, given by:

$$S(T) \sim \exp(X(T)).$$

Furthermore, for $Z \sim \mathcal{N}(\mu, \sigma^2)$, the following equalities hold,

$$\mathbb{E}\left[\mathrm{e}^Z\right] = \mathrm{e}^{\mu+\frac{1}{2}\sigma^2}, \quad \mathbb{V}\mathrm{ar}\left[\mathrm{e}^Z\right] = \left(\mathrm{e}^{\sigma^2} - 1\right)\mathrm{e}^{2\mu+\sigma^2}. \tag{2.16}$$

♦

Utilizing Itô's formula, we can thus determine the solution for $S(t)$ in (2.1), as

$$S(t) = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t - t_0) + \sigma(W^{\mathbb{P}}(t) - W^{\mathbb{P}}(t_0))\right). \tag{2.17}$$

For $X(t) := \log S(t)$, we see, based on the findings in Example 2.1.4, that

$$\begin{aligned} F_{X(t)}(x) \;&=\; \mathbb{P}[X(t) \leq x] = \\ &:= \frac{1}{\sigma\sqrt{2\pi(t-t_0)}} \int_{-\infty}^{x} \exp\left(-\frac{(z - \log S_0 - (\mu - \frac{\sigma^2}{2})(t-t_0))^2}{2\sigma^2(t-t_0)}\right)\mathrm{d}z, \end{aligned} \tag{2.18}$$

and correspondingly the probability density function, $f_X(x) := f_{X(t)}(x)$, reads

$$f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x}F_X(x) = \frac{1}{\sigma\sqrt{2\pi(t-t_0)}} \exp\left(-\frac{(x - \log S_0 - (\mu - \frac{\sigma^2}{2})(t-t_0))^2}{2\sigma^2(t-t_0)}\right). \tag{2.19}$$

The probability distribution function for $S(t)$ can now be obtained by the transformation $S(t) = \exp(X(t))$; for $x > 0$:

$$\begin{aligned} F_{S(t)}(x) &= \mathbb{P}[S(t) \leq x] = \mathbb{P}[\mathrm{e}^{X(t)} \leq x] = \mathbb{P}[X(t) \leq \log x] = F_{X(t)}(\log x) \\ &= \frac{1}{\sigma\sqrt{2\pi(t-t_0)}} \int_{-\infty}^{\log x} \exp\left(-\frac{(z - \log S_0 - (\mu - \frac{\sigma^2}{2})(t-t_0))^2}{2\sigma^2(t-t_0)}\right)\mathrm{d}z, \end{aligned}$$

which is lognormal. The probability density function for $S(t)$, $f_S(x) := f_{S(t)}(x)$, is found to be

$$f_S(x) = \frac{1}{\sigma x\sqrt{2\pi(t-t_0)}} \exp\left(-\frac{(\log \frac{x}{S_0} - (\mu - \frac{\sigma^2}{2})(t-t_0))^2}{2\sigma^2(t-t_0)}\right), \; x > 0. \tag{2.20}$$

In Figure 2.2 the time evolution of the probability density functions of the processes $X(t)$ and $S(t)$ are respectively presented. The parameter values are in the figure's caption. It can be seen from the graphs in Figure 2.2 that for a short time period the densities of $X(t)$ and $S(t)$ resemble each other well. This can be explained by the following relation between the normal and lognormal distributions. If $X \sim \mathcal{N}(0, 1)$, then $Y = \exp(X)$ is a lognormal distribution, and by the well-known MacLaurin expansion, see Equation (1.10), of the variable Y

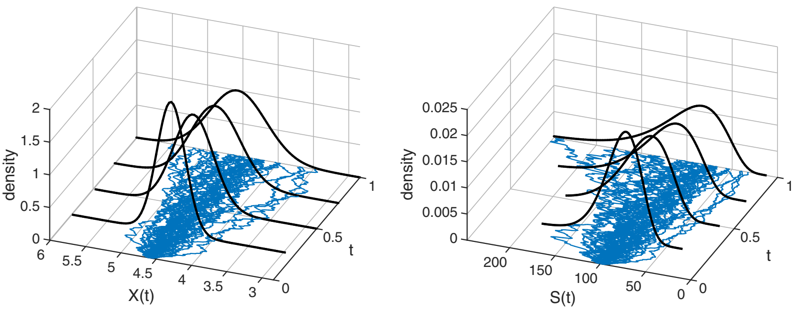


Figure 2.2: *Paths and the corresponding densities. Left:* $X(t) = \log S(t)$ *and Right:* $S(t)$ *with the following configuration:* $S_0 = 100$, $\mu = 0.05$, $\sigma = 0.4$; $T = 1$.

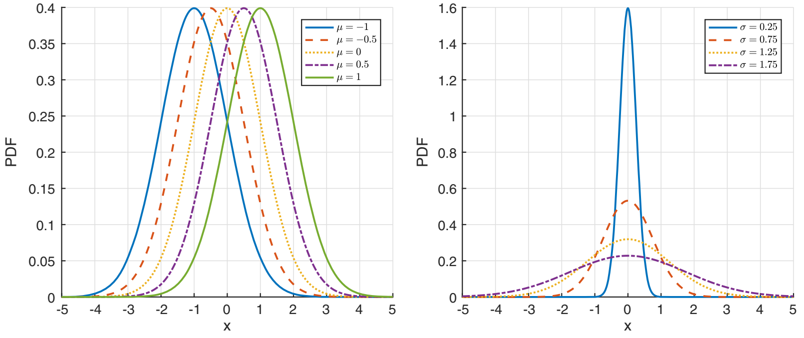


Figure 2.3: *Probability density function for the normal distribution. Left: the impact of parameter* μ, *with* $\sigma = 1$; *Right: the impact of* σ, *using* $\mu = 0$.

around 0, we find $Y \approx 1 + X \sim \mathcal{N}(1, 1)$. This approximation is insightful in the case when $T \to 0$.

Example 2.1.5 (Pictures of normal and lognormal distributions)
In Figures 2.3 and 2.4, graphical representations of the probability density functions, for different sets of parameters, are presented for respectively the normal and the lognormal distribution.

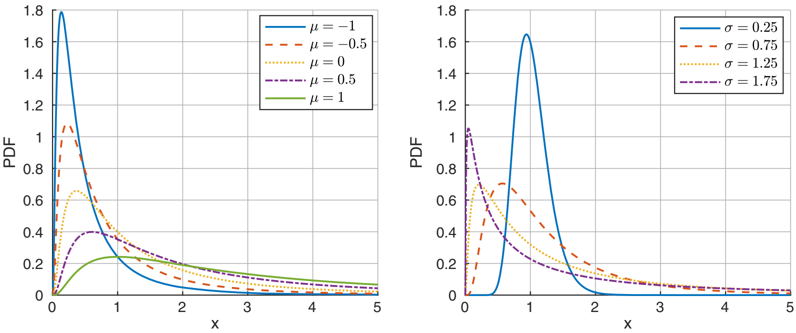


Figure 2.4: *Probability density function for the lognormal distribution. Left: the impact of parameter μ, with $\sigma = 1$; Right: impact of parameter σ, with $\mu = 0$.*

2.2 First generalizations

Here, we present some first variations of the basic GBM asset price process. We look into proportional dividends, which may be a first generalization of the GBM drift term, and into volatility generalizations, including a time-dependent volatility function.

2.2.1 Proportional dividend model

In the generic GBM stock price model in (2.1), a dividend payment has not been modeled. However, typically a company pays dividends once or twice a year. The exact amounts of dividend and the payment times may vary each year. At the time the dividend is paid, there will be a drop in the value of the stock.

Different mathematical models are available for the dividend payments, like deterministic or stochastic, continuous or discrete time models.

A continuous and constant dividend yield of size $q \cdot S(t)$, with a constant factor $q < 1$, is often used and it represents the fact that in a time instance $\mathrm{d}t$ the underlying asset pays out a *proportional dividend* of size $qS(t)\mathrm{d}t$. This is considered a suitable model particularly for stock indices (in which multiple stocks are represented). Arbitrage considerations indicate that the asset price must fall by the amount of dividend payment, i.e., we have in the case of proportional dividend payment,

$$\mathrm{d}S(t) = (\mu - q)S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{P}}(t), \tag{2.21}$$

modeling a dividend paying stock with a continuous stream of dividends.

2.2.2 Volatility variation

In the GBM model, the stock dynamics include the volatility parameter σ. One may wonder why the volatility is often written as σ, and why it is not set equal to $\sigma S(t)$, as this term $\sigma S(t)$ is in front of the random driver $\mathrm{d}W(t)$. When we use the word volatility, we actually mean "the volatility of the stock *return*". The stock performance over a time period $[t, t+\Delta t]$ is determined by the ratio, $\frac{S(t+\Delta t)-S(t)}{S(t)}$. For an instantaneously small period $\Delta t \to 0$ this ratio is given by

$$\lim_{\Delta t \to 0} \frac{S(t+\Delta t)-S(t)}{S(t)} = \frac{\mathrm{d}S(t)}{S(t)} = \mu \mathrm{d}t + \boxed{\sigma} \mathrm{d}W^{\mathbb{P}}(t). \tag{2.22}$$

From the stock return it is clear why σ is the volatility parameter of process $S(t)$.

Other asset models, in which the volatility in the $\mathrm{d}S(t)/S(t)$ term is written as a function of $S(t)$, are the so-called *parametric local volatility models*. Two well-known parametric local volatility models are the *quadratic model*, with its asset dynamics given by:

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu \mathrm{d}t + \boxed{\sigma S(t)} \mathrm{d}W^{\mathbb{P}}(t),$$

and the *so-called* 3/2 *model*, with

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu \mathrm{d}t + \boxed{\sigma \sqrt{S(t)}} \mathrm{d}W^{\mathbb{P}}(t).$$

These models form building blocks particularly for some stochastic interest rate processes. Some other parametric local volatility models will be discussed in follow-up sections.

2.2.3 Time-dependent volatility

Another generalization of the GBM asset price model, away from constant volatility, may be to use a time-dependent volatility coefficient, using $\sigma(t)$, i.e.

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma(t)S(t)\mathrm{d}W^{\mathbb{P}}(t), \quad S(t_0) = S_0, \tag{2.23}$$

where $\sigma(t)$ is a deterministic function. In the example below, we will match the moments of a model with constant volatility and one with time-dependent volatility.

Example 2.2.1 (Time-dependent volatility and moment matching)
As an illustrative example of moment matching, we consider the following two stochastic processes, $X(t)$ and $Y(t)$, that are governed by the following SDEs:

$$\mathrm{d}X(t) = \left(\mu - \frac{1}{2}\sigma^2(t)\right)\mathrm{d}t + \sigma(t)\mathrm{d}W^{\mathbb{P}}(t),$$

$$\mathrm{d}Y(t) = \left(\mu - \frac{1}{2}\sigma_*^2\right)\mathrm{d}t + \sigma_* \mathrm{d}W^{\mathbb{P}}(t),$$

with a time-dependent volatility $\sigma(t)$ for process $X(t)$ and a constant volatility parameter σ_* for process $Y(t)$. The expectations of $X(T)$ and $Y(T)$ can be determined as

$$\mathbb{E}[X(T)] = X_0 + \int_0^T \left(\mu - \frac{1}{2}\sigma^2(t)\right) \mathrm{d}t, \quad \mathbb{E}[Y(T)] = Y_0 + \left(\mu - \frac{1}{2}\sigma_*^2\right) T, \tag{2.24}$$

and the variance of process $X(t)$ reads:

$$\mathbb{V}\mathrm{ar}[X(T)] = \mathbb{E}[X^2(T)] - (\mathbb{E}[X(T)])^2 = \mathbb{E}\left[\int_0^T \sigma(t)\mathrm{d}W^{\mathbb{P}}(t)\right]^2 = \int_0^T \sigma^2(t)\mathrm{d}t,$$

where the last step is based on the Itô isometry (as in Section 1.3.2). The variance of process $Y(t)$ equals $\mathbb{V}\mathrm{ar}[Y(T)] = \sigma_*^2 T$. By equating the variances of the processes X and Y, i.e. $\mathbb{V}\mathrm{ar}[X(T)] = \mathbb{V}\mathrm{ar}[Y(T)]$, we find

$$\boxed{\sigma_* = \sqrt{\frac{1}{T}\int_0^T \sigma^2(t)\mathrm{d}t.}} \tag{2.25}$$

Using this particular choice of σ_*, also the expectations of $X(T)$ and $Y(T)$ (with $X_0 = Y_0$) are equal, as in (2.24) we will get

$$\begin{aligned}\mathbb{E}[Y(T)] &= Y_0 + \left(\mu - \frac{1}{2}(\sigma_*)^2\right) T \\ &= X_0 + \left(\mu - \frac{1}{2}\frac{1}{T}\int_0^T \sigma^2(t)\mathrm{d}t\right) T = \mathbb{E}[X(T)].\end{aligned}$$

The processes X and Y are log-transformed GBM asset prices and these are normally distributed. In that case only the first two moments have to match in order to guarantee equality in distribution. ♦

With the moment matching technique, both processes will have exactly the same marginal distributions, however, their transitional distributions will differ.

2.3 Martingales and asset prices

The asset price models discussed will satisfy the condition $\mathbb{E}[|S(t)|] < \infty$. Pricing of financial derivatives, like financial options, heavily depends on calculating the expectation of some payoff with respect to an underlying process. Therefore, it is crucial that the moments exist.

The martingale property implies that a mathematical model of a financial product is free of arbitrage, i.e. there are no risk-free profits present in the mathematical model for the stochastic quantity under consideration. The probability measure under which the martingale property holds is often called the *risk-neutral measure*, and it is denoted by $\mathbb{Q}$. Particularly, discounted tradable assets will be martingales under the risk-neutral measure.

2.3.1 $\mathbb{P}$-measure prices

Financial stock price processes as observed in the stock exchanges are not martingales usually (i.e. they are not completely unpredictable).

Example 2.3.1 (Expectation of $S(t)$ under $\mathbb{P}$-measure) From Equation (2.23) (the time-dependent volatility case) we have,

$$S(t) = S_0 \mathrm{e}^{\int_{t_0}^t (\mu - \frac{1}{2}\sigma^2(z))\mathrm{d}z + \int_{t_0}^t \sigma(z)\mathrm{d}W^{\mathbb{P}}(z)},$$

and the expectation is given by,

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[S(t)|\mathcal{F}(t_0)\right] &= \mathbb{E}^{\mathbb{P}}\left[S_0 \mathrm{e}^{\frac{1}{2}\int_{t_0}^t (\mu - \sigma^2(z))\mathrm{d}z + \int_{t_0}^t \sigma(z)\mathrm{d}W^{\mathbb{P}}(z)} \Big| \mathcal{F}(t_0)\right] \\ &= S_0 \mathrm{e}^{\mu(t-t_0) - \frac{1}{2}\int_{t_0}^t \sigma^2(z)\mathrm{d}z} \mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\int_{t_0}^t \sigma(z)\mathrm{d}W^{\mathbb{P}}(z)} \Big| \mathcal{F}(t_0)\right].\end{aligned}$$

For a normally distributed $X \sim \mathcal{N}(0,1)$, it can be derived that $\mathbb{E}[\mathrm{e}^{aX}] = \mathrm{e}^{\frac{1}{2}a^2}$, and,

$$\int_{t_0}^t \sigma(z)\mathrm{d}W^{\mathbb{P}}(z) \sim \mathcal{N}\left(0, \int_{t_0}^t \sigma^2(z)\mathrm{d}z\right).$$

Therefore,

$$\mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\int_{t_0}^t \sigma(z)\mathrm{d}W^{\mathbb{P}}(z)} \Big| \mathcal{F}(t_0)\right] = \mathrm{e}^{\frac{1}{2}\int_{t_0}^t \sigma^2(z)\mathrm{d}z},$$

so that,

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[S(t)|\mathcal{F}(t_0)\right] &= S_0 \mathrm{e}^{\mu(t-t_0) - \frac{1}{2}\int_{t_0}^t \sigma^2(z)\mathrm{d}z} \mathbb{E}^{\mathbb{P}}\left[\mathrm{e}^{\int_{t_0}^t \sigma(z)\mathrm{d}W^{\mathbb{P}}(z)} \Big| \mathcal{F}(t_0)\right] \\ &= S_0 \mathrm{e}^{\mu(t-t_0) - \frac{1}{2}\int_{t_0}^t \sigma^2(z)\mathrm{d}z} \mathrm{e}^{\frac{1}{2}\int_{t_0}^t \sigma^2(z)\mathrm{d}z} \\ &= S_0 \mathrm{e}^{\mu(t-t_0)}.\end{aligned}$$

♦

In a time interval Δt, we expect $S(t)$, as defined in Equation (2.1), to grow at some positive rate μ under the real-world probability measure $\mathbb{P}$. The stock price is governed by the lognormal distribution, with the expectation given by

$$\boxed{\mathbb{E}^{\mathbb{P}}[S(t)|\mathcal{F}(t_0)] = S_0 \mathrm{e}^{\mu(t-t_0)},}$$

so that $\mathbb{E}^{\mathbb{P}}[S(t+\Delta t)|\mathcal{F}(t)] > S(t)$, i.e. $S(t)$ is not a martingale.[2] In particular, we may assume that $\mu > r$, where r is the risk-free interest rate (as otherwise one would not invest in the asset), so that also for the discounted asset price process, we find

$$\mathbb{E}^{\mathbb{P}}[\mathrm{e}^{-r\Delta t} S(t+\Delta t)|\mathcal{F}(t)] > S(t).$$

[2] It is a so-called sub-martingale.

Let us therefore consider the stock process, $S(t)$, under a different measure, $\mathbb{Q}$, as follows

$$\mathrm{d}S(t) = rS(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{Q}}(t), \quad S(t_0) = S_0, \tag{2.26}$$

replacing the drift rate μ by the risk free interest rate r. The solution of Equation (2.26) reads:

$$S(t) = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t - t_0) + \sigma(W^{\mathbb{Q}}(t) - W^{\mathbb{Q}}(t_0))\right).$$

The concept of changing measures is very fundamental in computational finance. Chapter 7, particularly Section 7.2, is devoted to this concept, which is then used in almost all further chapters of this book.

2.3.2 $\mathbb{Q}$-measure prices

The stock price $S(t)$ is also governed by the lognormal distribution under this measure $\mathbb{Q}$, however, with the expectation given by:

$$\boxed{\mathbb{E}^{\mathbb{Q}}[S(t)|\mathcal{F}(t_0)] = S_0\mathrm{e}^{r(t-t_0)}.}$$

The expectation of $S(t)$, conditioned on time s with $s < t$, reads

$$\mathbb{E}^{\mathbb{Q}}[S(t)|\mathcal{F}(s)] = \mathrm{e}^{\log S(s) + r(t-s)} = S(s)\mathrm{e}^{r(t-s)} \neq S(s),$$

which implies that process $S(t)$ for $t \in [t_0, T]$ is also *not* a martingale under measure $\mathbb{Q}$.

With a *money-savings account*, i.e.,

$$M(t) = M(s)\mathrm{e}^{r(t-s)}, \tag{2.27}$$

we have

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\frac{S(t)}{M(t)}|\mathcal{F}(s)\right] &= \frac{\mathrm{e}^{-r(t-s)}}{M(s)}\mathbb{E}^{\mathbb{Q}}\left[S(t)|\mathcal{F}(s)\right] \\ &= \frac{\mathrm{e}^{-r(t-s)}}{M(s)}S(s)\mathrm{e}^{r(t-s)} = \frac{S(s)}{M(s)}.\end{aligned}$$

The compensation term in the definition of the discounted process is related to the *money-savings account* $M(t)$, i.e.

$$S(t)\mathrm{e}^{-r(t-t_0)} \stackrel{\text{def}}{=} \frac{S(t)}{M(t)},$$

where $\mathrm{d}M(t) = rM(t)\mathrm{d}t$, with $M(t_0) = 1$.

The money-savings account $M(t)$ is the so-called *numéraire*, expressing the unit of measure in which all other prices are expressed. Numéraire is a French word, meaning the basic standard by which values are measured.

So, we have a probability measure, $\mathbb{Q}$ (the risk-neutral measure) under which the stock price, discounted by the risk-free interest rate r, becomes a martingale, i.e.

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[e^{-r\Delta t}S(t+\Delta t)|\mathcal{F}(t)\right] &= \mathbb{E}^{\mathbb{Q}}\left[\frac{M(t)}{M(t+\Delta t)}S(t+\Delta t)|\mathcal{F}(t)\right]\\ &= M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{S(t+\Delta t)}{M(t+\Delta t)}|\mathcal{F}(t)\right]\\ &= M(t)\frac{S(t)}{M(t)} = S(t). \end{aligned} \tag{2.28}$$

The asset price which satisfies relation (2.28) is thus a GBM asset price process, in which the drift parameter is set equal to the risk-free interest rate $\mu = r$, compare (2.1) and (2.26).

Example 2.3.2 (Asset paths) In Figure 2.5 we present some discrete paths that have been generated by a GBM process with $\mu = 0.15$, $r = 0.05$ and $\sigma = 0.1$. The left-hand picture in Figure 2.5 displays the paths of a discounted stock process, $S(t)/M(t)$, where $S(t)$ is defined under the real-world measure $\mathbb{P}$, where it has a drift parameter $\mu = 0.15$. The right-hand picture shows the paths for $S(t)/M(t)$, where $S(t)$ has drift parameter $r = 0.05$. Since $\mu > r$, the discounted stock price under the real-world measure will increase much faster than the stock process under the risk-neutral measure. In other words, under the real-world measure $S(t)/M(t)$ is a sub-martingale.

As we will see later, the valuation of financial options is related to the computation of expectations. The notations $\mathbb{E}^{\mathbb{Q}}[g(T,S)]$ and $\mathbb{E}^{\mathbb{P}}[g(T,S)]$ indicate that we take the expectation of a certain function, $g(T,S)$, which depends on the stochastic asset price process $S(t)$. Asset price $S(t)$ can be defined under a specific measure, like under $\mathbb{P}$ or $\mathbb{Q}$, where $\mathbb{P}[\cdot]$ and $\mathbb{Q}[\cdot]$ indicate the associated probabilities. If we take, for example, $g(T,S) = \mathbb{1}_{S(T)>0}$, then the following equalities hold:

$$\mathbb{E}^{\mathbb{P}}[\mathbb{1}_{S(T)>0}] = \int_\Omega \mathbb{1}_{S(T)>0}\mathrm{d}\mathbb{P} = \mathbb{P}[S(T)>0], \tag{2.29}$$

$$\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{S(T)>0}] = \int_\Omega \mathbb{1}_{S(T)>0}\mathrm{d}\mathbb{Q} = \mathbb{Q}[S(T)>0], \tag{2.30}$$

where the notation $\mathbb{1}_{X\in\Omega}$ stands for the indicator function of the set Ω, as in (1.1).

Superscripts $\mathbb{Q}$ and $\mathbb{P}$ indicate which process is used in the expectation operator to calculate the corresponding integrals. In other words, it indicates the underlying dynamics that have been used.

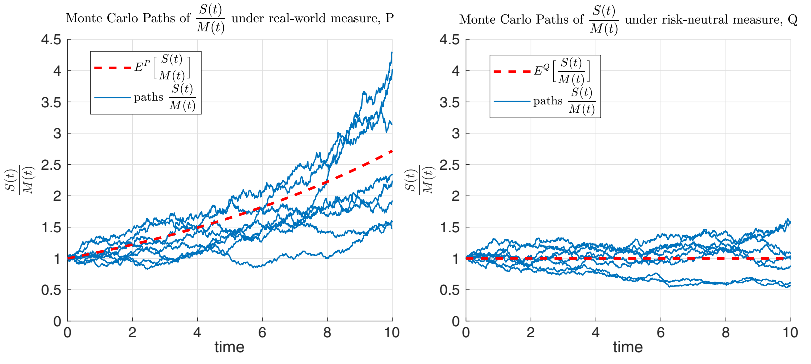


Figure 2.5: *Paths for the discounted stock process, $S(t)/M(t)$, with $S(t)$ under real-world measure $\mathbb{P}$ (left) and $S(t)$ under risk-neutral measure $\mathbb{Q}$ (right), with parameters $r = 0.05$, $\mu = 0.15$ and $\sigma = 0.1$.*

2.3.3 Parameter estimation under real-world measure $\mathbb{P}$

In this section we give insight in the parameter estimation for a stochastic process under the real-world measure $\mathbb{P}$. Estimation under the real-world measure means that the model parameters for a stochastic process are obtained by a *calibration to historical stock price values.* Calibration will be completely different under the risk-neutral measure $\mathbb{Q}$, as we will discuss in the chapter to follow. The parameters here will be estimated by using a very popular statistical estimation technique, which is called "maximum likelihood estimation method" (MLE) [Harris and Stocker, 1998]. The idea behind the method is to find the parameter estimates of the underlying probability distribution for a given data set.

As a first example of a stochastic process let us take the arithmetic Brownian motion process, which was introduced by Louis Bachelier [Bachelier, 1900] and is given by:

$$\mathrm{d}X(t) = \mu \mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{P}}(t), \tag{2.31}$$

with its solution:

$$\begin{aligned} X(t) &= X(t_0) + \mu(t - t_0) + \sigma(W^{\mathbb{P}}(t) - W^{\mathbb{P}}(t_0)), \quad \text{or} \\ X(t) &\sim \mathcal{N}\left(X(t_0) + \mu(t - t_0), \sigma^2(t - t_0)\right). \end{aligned} \tag{2.32}$$

In order to forecast "tomorrow's" (at time $t + \Delta t$) value, given that we observed the values up to now (where "now" is time t), we can calculate the conditional expectation as follows,

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left[X(t+\Delta t)|\mathcal{F}(t)\right] &= \mathbb{E}^{\mathbb{P}}\left[X(t) + \mu(t + \Delta t - t) + \sigma(W^{\mathbb{P}}(t+\Delta t) - W^{\mathbb{P}}(t))|\mathcal{F}(t)\right] \\ &= \mu \Delta t + \mathbb{E}^{\mathbb{P}}\left[X(t)|\mathcal{F}(t)\right] + \mathbb{E}^{\mathbb{P}}\left[\sigma(W^{\mathbb{P}}(t+\Delta t) - W^{\mathbb{P}}(t))|\mathcal{F}(t)\right]. \end{aligned}$$

Because of the measurability principle, $\mathbb{E}^{\mathbb{P}}\left[X(t)|\mathcal{F}(t)\right] = X(t)$ and by the property of independent increments the second expectation is equal to zero, i.e.,

$$\mathbb{E}^{\mathbb{P}}\left[X(t+\Delta t)|\mathcal{F}(t)\right] = X(t) + \mu\Delta t.$$

For the variance, we obtain,

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[X(t+\Delta t)|\mathcal{F}(t)\right] \\
&\quad = \mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[X(t) + \mu(t+\Delta t - t) + \sigma(W^{\mathbb{P}}(t+\Delta t) - W^{\mathbb{P}}(t))|\mathcal{F}(t)\right] \\
&\quad = \mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[X(t)|\mathcal{F}(t)\right] + \mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[\sigma\sqrt{\Delta t}\widetilde{Z}|\mathcal{F}(t)\right],
\end{aligned}$$

with $\widetilde{Z} \sim \mathcal{N}(0,1)$. By Equation (2.32) and the fact that[3] $\mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[X(t)|\mathcal{F}(t)\right] = 0$, we find:

$$\mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[X(t+\Delta t)|\mathcal{F}(t)\right] = \sigma^2\Delta t.$$

So, we arrive at the prediction of the $t+\Delta t$ value, given the information[4] at time t, i.e. at the conditional random variable,

$$X(t+\Delta t)|X(t) \sim \mathcal{N}\left(X(t) + \mu\Delta t, \sigma^2\Delta t\right), \tag{2.33}$$

which is a normal random variable for which the parameters μ and σ have to be estimated.

Parameter estimation for arithmetic Brownian motion

Since the conditional distribution is normally distributed, the log-likelihood functional form is known. Assuming independence of the observations and by setting the observation values (typically historical values of some return or log-asset) $X(t_0), X(t_1), \ldots, X(t_m)$, the likelihood function $L(\hat{\mu}, \hat{\sigma}^2|X(t_0), X(t_1), \ldots, X(t_m))$ is given by,

$$L(\hat{\mu}, \hat{\sigma}^2|X(t_0), X(t_1), \ldots, X(t_m)) = \prod_{k=0}^{m-1} f_{X(t_{k+1})|X(t_k)}(X(t_{k+1})), \tag{2.34}$$

with $\hat{\mu}$ and $\hat{\sigma}$ to be estimated, By Equation (2.33) we find:

$$f_{X(t_{k+1})|X(t_k)}(x) = f_{\mathcal{N}(X(t_k)+\hat{\mu}\Delta t, \hat{\sigma}^2\Delta t)}(x), \tag{2.35}$$

---

[3]Note that because of the measurability, we find:

$$\mathbb{V}\mathrm{ar}^{\mathbb{P}}\left[X(t)|\mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{P}}\left[X(t)^2|\mathcal{F}(t)\right] - \left(\mathbb{E}^{\mathbb{P}}\left[X(t)|\mathcal{F}(t)\right]\right)^2 = X^2(t) - X^2(t) = 0.$$

[4]By the notation $|X(t)$, we also indicate "the knowledge up-to time t", like when using $\mathcal{F}(t)$.

and Equation (2.34) reads,

$$\begin{aligned}
&L(\hat{\mu}, \hat{\sigma}^2 | X(t_0), X(t_1), \ldots, X(t_m)) \\
&\quad = \prod_{k=0}^{m-1} f_{\mathcal{N}(X(t_k)+\hat{\mu}\Delta t, \hat{\sigma}^2\Delta t)}(X(t_{k+1})) \\
&\quad = \prod_{k=0}^{m-1} \frac{1}{\sqrt{2\pi\hat{\sigma}^2\Delta t}} \exp\left(-\frac{(X(t_{k+1}) - X(t_k) - \hat{\mu}\Delta t)^2}{2\hat{\sigma}^2\Delta t}\right).
\end{aligned}$$

To simplify the maximization, we use the logarithmic transformation, i.e.,

$$\log L(\hat{\mu}, \hat{\sigma}^2 | \ldots) = \log\left(\frac{1}{\sqrt{2\pi\hat{\sigma}^2\Delta t}}\right)^m - \sum_{k=0}^{m-1} \frac{(X(t_{k+1}) - X(t_k) - \hat{\mu}\Delta t)^2}{2\hat{\sigma}^2\Delta t}.$$

Now,

$$\log\left(\frac{1}{\sqrt{2\pi\hat{\sigma}^2\Delta t}}\right)^m = \log\left(2\pi\hat{\sigma}^2\Delta t\right)^{-0.5m} = -0.5m\log\left(2\pi\hat{\sigma}^2\Delta t\right),$$

by which we obtain,

$$\log L(\hat{\mu}, \hat{\sigma}^2 | \ldots) = -0.5m\log\left(2\pi\hat{\sigma}^2\Delta t\right) - \sum_{k=0}^{m-1} \frac{(X(t_{k+1}) - X(t_k) - \hat{\mu}\Delta t)^2}{2\hat{\sigma}^2\Delta t}.$$

To determine the maximum of the log-likelihood, the first-order conditions have to be satisfied for both parameters,

$$\frac{\partial}{\partial\hat{\mu}} \log L(\hat{\mu}, \hat{\sigma}^2 | \ldots) = 0, \qquad \frac{\partial}{\partial\hat{\sigma}^2} \log L(\hat{\mu}, \hat{\sigma}^2 | \ldots) = 0.$$

This gives us the following estimators:

$$\hat{\mu} = \frac{1}{m\Delta t}(X(t_m) - X(t_0)), \quad \hat{\sigma}^2 = \frac{1}{m\Delta t}\sum_{k=0}^{m-1}(X(t_{k+1}) - X(t_k) - \hat{\mu}\Delta t)^2. \tag{2.36}$$

Based on these estimators $\hat{\mu}$, $\hat{\sigma}^2$, for given historical values $X(t_0), X(t_1), \ldots, X(t_m)$, we may determine the "historical" parameters for the process $X(t)$.

Log-normal distribution

Using the methodology above, in this numerical experiment we will forecast the stock value of an electric vehicle company. The ABM process we discussed earlier is not the most adequate process for this purpose, as it may give rise to negative stock values. An alternative is GBM, which may be closely related to ABM. Under the real-world measure $\mathbb{P}$ the following process for the stock is considered,

$$\mathrm{d}S(t) = \left(\mu + \frac{1}{2}\sigma^2\right) S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{P}}(t). \tag{2.37}$$

By the log-transformation and using Itô's lemma, we find,

$$\mathrm{d}X(t) := \mathrm{d}\log S(t) = \mu\mathrm{d}t + \sigma\mathrm{d}W^{\mathbb{P}}(t). \tag{2.38}$$

So, under the log-transformation of the process $S(t)$ in (2.37), we arrive at the estimates that were obtained from ABM in (2.36).

Now, given the observation of electric vehicle company stock prices $S(t_i)$ (in Figure 2.6), first we perform the log-transformation to obtain $X(t) := \log S(t)$ and subsequently estimate the parameters $\hat{\mu}$ and $\hat{\sigma}$ in (2.36). The data set contains closing prices between 2010 and 2018, giving us the following estimators $\hat{\mu} = 0.0014$ and $\hat{\sigma} = 0.0023$.

With $\hat{\mu}$ and $\hat{\sigma}$ for μ, σ in (2.37) determined, we may simulate "future" realizations for the stock prices. In Figure 2.6 (left) the historical stock prices are presented, whereas in the right-side figure the forecast for the future values is shown. In the experiments this forecast of the future prices is purely based on the historical stock realizations. This is the essence of working under the $\mathbb{P}$-real world measure. Note that by assuming a GBM process the stock prices are expected to grow in time. However, by changing the assumed stochastic process we may obtain a different forecast.

Estimating the parameter μ under measure $\mathbb{P}$ is thus related to fitting the process in (2.1) to *historical stock values*. Note that the paths in Figure 2.5 suggest that under the real-world measure the expected returns of the discounted

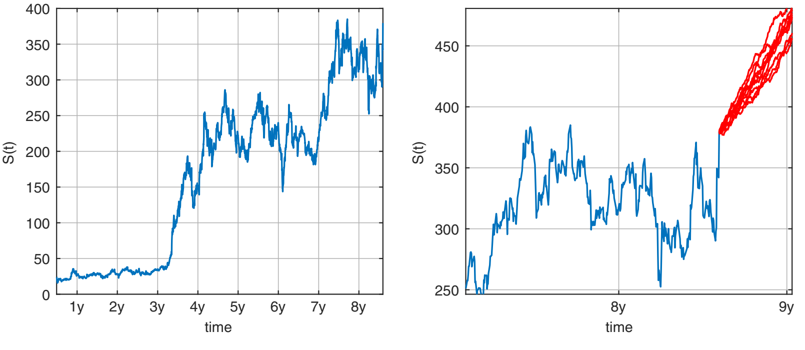


Figure 2.6: *Left: historical stock values of electric car company; Right: forecast for the future performance of the stock.*

stock are typically higher than one would expect under the risk-neutral measure. This is partially an argument why certain hedge funds work with processes under the $\mathbb{P}$-measure, where they aim to forecast the stock prices using Econometrics tools. Speculation is also typically done under this measure. Risk management, meaning analyzing the future behavior of assets in the real world, as it is often done by the regulator of financial institutions, is usually done under measure $\mathbb{P}$ as well. In this case, specific asset scenarios (*back testing*) or even when *stress-testing*, extreme asset scenarios are provided under which the companies' balance sheets should be evaluated.

When we deal with financial option valuation in the chapters to follow, however, we need to fit the parameter values under the so-called *risk-neutral measure* $\mathbb{Q}$. Then, we are mainly interested in the parameters *for a future time period*, and wish to extract the relevant information from financial products that may give us information about the expectations of the market participants about the future performance of the financial asset.

This means that we aim to extract the relevant *implied asset information* from arbitrage-free option contracts and other financial derivatives. Financial institutions work under measure $\mathbb{Q}$ when pricing financial derivative products for their clients.

Pension funds should worry about $\mathbb{P}$ measure valuation (liabilities and their performance and risk management in the real world), and also about $\mathbb{Q}$ measure calculations, when derivative contracts form a part of the pension investment portfolios. In essence, hedge funds may "bet" on the future, while banks and pension funds get a premium upfront and they hedge their position to keep the premium intact.

2.4 Exercise set

Exercise 2.1 Apply Itô's lemma to find:

a. The dynamics of process $g(t) = S^2(t)$, where $S(t)$ follows a log-normal Brownian motion given by:
$$\frac{\mathrm{d}S(t)}{S(t)} = \mu \mathrm{d}t + \sigma \mathrm{d}W(t),$$
with constant parameters μ, σ and Wiener process $W(t)$.

b. The dynamics for $g(t) = 2^{W(t)}$, where $W(t)$ is a standard Brownian motion. Is this a martingale?

Exercise 2.2 Apply Itô's lemma to show that:

a. $X(t) = \exp(W(t) - \frac{1}{2}t)$ solves $\mathrm{d}X(t) = X(t)\mathrm{d}W(t)$,

b. $X(t) = \exp(2W(t) - t)$ solves $\mathrm{d}X(t) = X(t)\mathrm{d}t + 2X(t)\mathrm{d}W(t)$,

Exercise 2.3 Derive the Itô integration-by-parts rule, which reads,

$$\int_0^T \mathrm{d}X(t)\mathrm{d}Y(t) = (X(t)Y(t))|_{t=0}^{t=T} - \int_0^T X(t)\mathrm{d}Y(t) - \int_0^T Y(t)\mathrm{d}X(t). \tag{2.39}$$

This can be written in differential form, as

$$\mathrm{d}(X(t) \cdot Y(t)) = Y(t) \cdot \mathrm{d}X(t) + X(t) \cdot \mathrm{d}Y(t) + \mathrm{d}X(t) \cdot \mathrm{d}Y(t). \tag{2.40}$$

where the additional term, $\mathrm{d}X(t)\mathrm{d}Y(t)$ does not appear in the deterministic integration-by-parts rule.

Derive this rule by means of the following discrete sum,

$$\sum_{k=1}^{N} (X(t_{k+1}) - X(t_k))(Y(t_{i+k}) - Y(t_k)).$$

Exercise 2.4 For this exercise, it is necessary to download some freely available stock prices from the web.

a. Find two data sets with daily stock prices of two different stocks S_1 and S_2, that are "as independent as possible". Check the independence by means of a scatter plot of daily returns.

b. Find yourself two data sets with daily stock prices of two different stocks S_3 and S_4, that are "as dependent as possible". Check the dependence by means of a scatter plot of daily returns.

Exercise 2.5 Asset price process $S(t)$ is governed by the geometric Brownian motion,

$$S(t) = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma\sqrt{t}Z}, \text{ where } Z \sim \mathcal{N}(0,1)$$

with the lognormal density function,

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi t}} \exp\left(\frac{-(\log(x/S_0) - (\mu - \sigma^2/2)t)^2}{2\sigma^2 t}\right), \text{ for } x > 0.$$

Determine $\mathbb{E}[S(t)]$ and $\mathbb{V}\mathrm{ar}[S(t)]$.

Exercise 2.6 Choose two realistic values: $0.1 \leq \sigma \leq 0.75$, $0.01 \leq \mu \leq 0.1$.

a. With $T = 3$, $S_0 = 0.7$, $\Delta t = 10^{-2}$ generate 10 asset paths that are driven by a geometric Brownian motion and the parameters above.

b. Plot for these paths the "running sum of square increments", i.e.

$$\sum_{k=1}^{m} (\Delta S(t_k))^2 .$$

c. Use asset price market data (those from Exercise 2.4) and plot the asset return path and the running sum of square increments.

Exercise 2.7 Consider a stock price process, $S(t) = \exp(X(t))$, and determine the relation between the densities of $S(t)$ and $X(t)$. *Hint:* show that

$$f_{S(t)}(x) = \frac{1}{x} f_{X(t)}\big(\log(x)\big), \quad x > 0.$$

Exercise 2.8 Use Itô's lemma to prove that,

$$\int_0^T W^2(t)\mathrm{d}W(t) = \frac{1}{3}W^3(T) - \int_0^T W(t)\mathrm{d}t.$$

Exercise 2.9 Suppose that $X(t)$ satisfies the following SDE,

$$\mathrm{d}X(t) = 0.04X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}W^{\mathbb{P}}(t),$$

and $Y(t)$ satisfies:

$$\mathrm{d}Y(t) = \beta Y(t)\mathrm{d}t + 0.15Y(t)\mathrm{d}W^{\mathbb{P}}(t).$$

Parameters β, σ are positive constants and both processes are driven by the same Wiener process $W^{\mathbb{P}}(t)$.
For a given process

$$Z(t) = 2\frac{X(t)}{Y(t)} - \lambda t,$$

with $\lambda \in \mathbb{R}^+$.

a. Find the dynamics for $Z(t)$.

b. For which values of β and λ is process $Z(t)$ a martingale?

CHAPTER 3

The Black-Scholes Option Pricing Equation

In this chapter:

Financial derivatives are products that are based on the performance of some underlying asset, like a stock, an interest rate, an foreign-exchange rate, or a commodity price. They are introduced in this chapter, in **Section 3.1**. The fundamental pricing partial differential equation for the valuation of European options is derived in that section. For options on stocks, it is the famous *Black-Scholes equation*.
In this chapter, we also discuss a solution approach for the option pricing partial differential equation, which is based on the *Feynman-Kac theorem*, in **Section 3.2**. This theorem connects the solution of the Black-Scholes equation to the calculation of the discounted expected payoff function, under the risk-neutral measure. This formulation of the option price gives us a semi-analytic solution for the Black-Scholes equation. A hedging experiment is subsequently described in **Section 3.3**.

Keywords: Black-Scholes partial differential equation for option valuation, Itô's lemma, Feynman-Kac theorem, discounted expected payoff approach, martingale approach, hedging.

3.1 Option contract definitions

An option contract is a financial contract that gives the holder the right to trade (buy or sell) in an underlying asset in the future at a predetermined price. In fact, the option contract offers its holder the *optionality* to buy or sell the asset; there is no obligation.

If the underlying asset does not perform favorably, the holder of the option does not have to exercise the option and thus does not have to trade in the asset.

The counterparty of the option contract, the option seller (also called the *option writer*), is however obliged to trade in the asset when the holder makes use of the exercise right in the option contract. Options are also called *financial derivatives*, as their values are derived from the performance of another underlying asset. There are many different types of option contracts. Since 1973, standardized option contracts have been traded on regulated exchanges. Other option contracts are directly sold by financial companies to their customers.

3.1.1 Option basics

Regarding financial option contracts, there is the general distinction between call and put options.

A *call option* gives an option holder the right to *purchase an asset*, whereas a *put option* gives a holder the right to *sell an asset*, at some time in the future $t = T$, for a prescribed amount, *the strike price* denoted by K.

In the case of the so-called *European option*, there is one prescribed time point in the future in the option contract, which is called *the maturity time or expiry date* (denoted by $t = T$), at which the holder of the option may decide to trade in the asset for the strike price.

Example 3.1.1 (Call option example) Let's consider a specific scenario for a call option holder who bought, at $t = t_0 = 0$, a call option with expiry date $T = 1$ and strike price $K = 150$ and where the initial asset price $S_0 = 100$, see Figure 3.1 (left side). Suppose we are now at time point $t = t_0$. In the left picture of Figure 3.1 two possible future asset price path scenarios are drawn. When the blue asset path would occur, the holder of the call option would exercise the option at time $t = T$. At that time, the option holder would pay K to receive the asset worth $S(T) = 250$. Selling the asset immediately in the market would give the holder the positive amount $250 - 150 = 100$.

In the case the red colored asset path would happen, however, the call option would not be exercised by the holder as the asset would be worth less than K in the financial market. As the call option contract has the optionality to exercise, there is no need to exercise when $S(T) < K$ (why buy something for a price K when it is cheaper in the market?). ♦

Generally, the holder of a European call option will logically exercise the call option, when the asset price at maturity time $t = T$ is larger than the strike price, i.e. when $S(T) > K$. In the case of exercise, the option holder will thus pay an amount K to the writer, to obtain an asset worth $S(T)$. The profit in this case is $S(T) - K$, as the asset may immediately be sold in the financial market.

If $S(T) < K$, the holder will *not* exercise the European call option and the option is in that case valueless. In that case, the option holder would be able to buy the asset for less than the strike price K in the financial market, so it does not make any sense to make use of the right in the option contract.

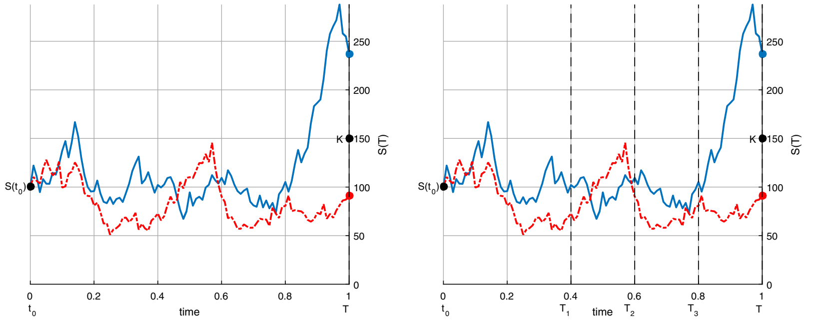


Figure 3.1: *Sketch of stochastic behavior of an asset; left: representation of a contract with only one decision point; right: a contract with more than one decision point, T_i, $i = 1, \ldots$.*

With the payoff function denoted by $H(T, S)$, we thus find for a call option, see Figure 3.2 (left side picture),

$$V_c(T, S) = H(T, S) := \max(S(T) - K, 0), \tag{3.1}$$

where the value of the call option at time t, for certain asset value $S = S(t)$ is denoted by $V_c(t, S)$, and $H(T, \cdot)$ is the option's payoff function.

The put option will be denoted by $V_p(t, S)$. Without any specification of the call or put features, the option is generally written as $V(t, S)$. However, we also sometimes use the notation $V(t, S; K, T)$, to emphasize the dependence on strike price K and maturity time T, when this appears useful. In the chapters dealing with options on interest rates we simply use the notation $V(t)$, because the option value may depend on many arguments in that case (too many to write down).

A European put option gives the holder the right to *sell* an asset at the maturity time T for a strike price K. The writer is then obliged to buy the asset, if the holder decides to sell.

At the maturity time, $t = T$, a European put option has the value $K - S(T)$, if $S(T) < K$. The put option expires worthless if $S(T) > K$, see Figure 3.2 (right side), i.e.,

$$V_p(T, S(T)) = H(T, S(T)) := \max(K - S(T), 0), \tag{3.2}$$

where a put option is denoted by $V_p(t, S(t))$.

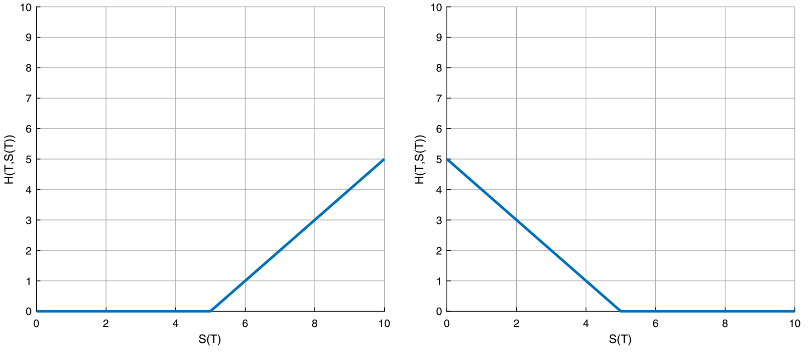


Figure 3.2: *The payoff $H(T, S(T))$ of European options with strike price K, at maturity time T; left: for a call option, right: for a put option.*

Definition 3.1.1 *The put and call payoff functions are convex functions.*
Remember that function $g(x)$ is called convex *if,*

$$g(\theta x_1 + (1-\theta)x_2) \leq \theta \cdot g(x_1) + (1-\theta) \cdot g(x_2), \; \forall x_1, x_2 \in \mathcal{C}, \forall \theta \in [0,1],$$

with $\mathcal{C}$ a convex set in $\mathbb{R}$. We have,

$$\max(\theta x_1 + (1-\theta)x_2, 0) \leq \theta \cdot \max(x_1, 0) + (1-\theta) \cdot \max(x_2, 0), \; \forall x_1, x_2 \in \mathcal{C}, \forall \theta \in [0,1].$$

The max *functions in the payoffs, are convex operators, because the functions 0, $x - K$ and $K - x$ are all convex.* ◀

Definition 3.1.2 (ITM, OTM, ATM) *When for an underlying asset value at a time point $t < T$, i.e. $S(t)$, the corresponding option's payoff would be equal to zero (like* $\max(K - S(t), 0) = 0$*, for a put), we say that the* option is "out-of-the money" (OTM)*. Its* intrinsic option value *($K - S(t)$ for a put) then equals zero. When the option's intrinsic value is positive, the option is said to be* in-the-money (ITM)*. When the intrinsic value is close to zero (like $K - S(t) \approx 0$, for a put), we call the option* at-the-money (ATM)*.*
The reasoning for puts and calls goes similarly, with the appropriate payoff functions. ◀

Some other financial derivative contracts include so-called *early exercise features*, meaning that there is *more than just one date, $t = T$,* to exercise these options. In the case of American call options, for example, the exercise of the option, i.e. paying K to purchase an asset worth $S(t)$, is permitted *at any time* during the life of the option, $t_0 \leq t \leq T$, whereas Bermudan options can be exercised *at certain dates, T_i, $i = 1, \ldots m$,* until the maturity time $T_m = T$ (see the right side picture of Figure 3.1).

An American option is thus a *continuous time equivalent* of the *discrete time* Bermudan option, assuming an increasing number of exercise dates, $m \to \infty$, with increasingly smaller time intervals, $\Delta t = \frac{T}{m} \to 0$. Options on individual stocks traded at regulated exchanges are typically American options.

Another class of options is defined as the *exotic options*, as they are governed by so-called exotic features in their payoff functions. These features can be path-dependency aspects of the stock prices, where the payoff does not only depend on the stock price $S(t)$ or $S(T)$, but also on certain functions of the stock price at different time points. These options are not traded on regulated exchanges but are sold *over-the-counter* (OTC), meaning that they are sold directly by banks and other financial companies to their counterparties. We will encounter some of these contracts, and the corresponding efficient pricing techniques, in the chapters to follow. Excellent overviews on the derivation of the governing PDEs in finance, for a variety of options, like exotic options and American options, include [Wilmott, 1998; Wilmott *et al.*, 1995; Hull, 2012].

Put-call parity

The put-call parity for European options represents a fundamental result in quantitative finance. It is based on the following reasoning. Suppose we have two portfolios, one with a European put option and a stock, i.e., $\Pi_1(t, S) = V_p(t, S) + S(t)$, and a second portfolio with a European call option bought and the present value of cash amount K, i.e. $\Pi_2(t, S) = V_c(t, S) + K\mathrm{e}^{-r(T-t)}$. The term $\mathrm{e}^{-r(T-t)}$ represents the discount factor, i.e., the value of €1 paying at time T as seen from today, with a continuously compounding interest rate r.

At maturity time T, we find

$$\Pi_1(T, S) = \max(K - S(T), 0) + S(T),$$

which equals K, if $S(T) < K$, and has value $S(T)$, if $S(T) > K$. In other words, $\Pi_1(T, S) = \max(K, S(T))$. The second portfolio has exactly the same value at $t = T$, i.e. $\Pi_2(T, S) = \max(K, S(T))$.

If $\Pi_1(T, S) = \Pi_2(T, S)$, then this equality should also be satisfied *at any time prior to the maturity time* T. Otherwise, an obvious arbitrage would occur, consisting of buying the cheaper portfolio and selling the more expensive one, so that at $t < T$ a positive cash amount results from this strategy. At $t = T$, both portfolios will have the same value so that buying one of the portfolios and selling the other at $t = T$ will result in neither profits nor losses at $t = T$, and the profit achieved at $t < T$ would remain. This is an arbitrage, i.e., a risk-free profit is obtained which is greater than just putting money on a bank account. This cannot be permitted. For any $t < T$, therefore the put-call parity

$$\boxed{V_c(t, S) = V_p(t, S) + S(t) - K\mathrm{e}^{-r(T-t)},} \tag{3.3}$$

should be satisfied. The interest rate r is supposed to be constant here.

In the case of a *dividend paying stock*, with a continuous stream of dividends, a put-call parity is also valid, but it should be modified. For the stock price $S(t)$

a dividend stream should be taken into account, i.e. $S(t)\mathrm{e}^{-q(T-t)}$, leading to the following adapted put-call parity relation,

$$\boxed{V_c(t,S) = V_p(t,S) + S(t)\mathrm{e}^{-q(T-t)} - K\mathrm{e}^{-r(T-t)}.} \tag{3.4}$$

3.1.2 Derivation of the partial differential equation

An important question in quantitative finance is what is a fair value for a financial option at the time of selling the option, i.e. at $t = t_0$. In other words, how to determine $V(t_0, S_0)$, or, more generally, $V(t,S)$ for any $t \geq t_0$. Another question is how an option writer can reduce the risk of trading in the asset $S(T)$ at time $t = T$, for a fixed strike price K. In other words, how to manage the risk of the option writer?

Based on the assumption of a geometric Brownian Motion process for the asset prices $S(t)$, Fisher Black and Myron Scholes derived their famous partial differential equation (PDE) for the valuation of European options, published in 1973 in the Journal of Political Economy [Black and Scholes, 1973]. The Black-Scholes model is one of the most important models in financial derivative pricing.

The derivation to follow is based on the assumption that the interest rate r and volatility σ are constants or known functions of time. Further, a liquid financial market is assumed, meaning that assets can be bought and sold any time in arbitrary amounts. Short-selling is allowed, so that negative amounts of the asset can be traded as well, and transaction costs are not included (neither is a dividend payment during the lifetime of the option). In the model, there is no bid-ask spread for the stock prices nor for the option prices.

In this section, we will present the main ideas by Fisher Black and Myron Scholes. The derivation of the pricing PDE is based on the concept of a *replicating portfolio*, which is updated each time step. This is a portfolio, which is set up by the seller of the option and has essentially the same cash flows (payments) as the option contract which is sold by the option writer. A replicating portfolio can be either a static or a dynamic portfolio — it depends on whether one needs to update (re-balance) the financial position in time or not. A static portfolio is set up once, after selling the financial contract, and the portfolio is then not changed during the lifetime of the contract. A dynamic portfolio is regularly updated, based on the available newly updated market information. We follow the so-called dynamic delta hedge strategy with a dynamically changing (re-balanced) replicating portfolio.

Let's start with a stock price process, $S \equiv S(t)$, which is the underlying for the financial derivative contract, $V \equiv V(t,S)$, representing the value of a European option (sometimes also called "a *plain vanilla contingent claim*"). The underlying stock price process is assumed to be GBM, with the dynamics under the real-world measure $\mathbb{P}$ as in (2.1), i.e.,

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W^{\mathbb{P}}(t).$$

As the price of the option $V(t,S)$ is a function of time t and the stochastic process $S(t)$, we will derive its dynamics, with the help of Itô's lemma, as follows:

$$\begin{aligned} \mathrm{d}V(t,S) &= \frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial V}{\partial S}\mathrm{d}S + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\mathrm{d}S)^2 \\ &= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t + \sigma S \frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{P}}. \end{aligned} \tag{3.5}$$

We construct a *portfolio* $\Pi(t,S)$, consisting of one option with value $V(t,S)$ and an amount, $-\Delta$, of stocks with value $S(t)$,

$$\boxed{\Pi(t,S) = V(t,S) - \Delta S(t).} \tag{3.6}$$

This portfolio thus consists of one long position in the option $V(t,S)$, and a short position of size Δ in the underlying $S(t)$.

Remark 3.1.1 (Short-selling of stocks) *When a trader holds a* negative number *of stocks, this implies that she/he has been short-selling stocks. Intuitively, it is quite confusing to own a negative amount of something. If you borrow money, then, in principle, you own a negative amount of money, because the money has to be paid back at some time in the future.*

Similarly, in financial practice, short-selling means that the trader borrows stocks from a broker, at some time t, for a fee. The stocks are subsequently sold, but the trader has to return the stocks to the broker at some later time. ▲

By Itô's lemma we derive the dynamics for an infinitesimal change in portfolio value $\Pi(t,S)$, based on the asset dynamics from Equation (2.1):

$$\boxed{\begin{aligned} \mathrm{d}\Pi &= \mathrm{d}V - \Delta \mathrm{d}S \\ &= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t + \sigma S \frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{P}} - \Delta\left[\mu S\mathrm{d}t + \sigma S \mathrm{d}W^{\mathbb{P}}\right] \\ &= \left[\frac{\partial V}{\partial t} + \mu S\left(\frac{\partial V}{\partial S} - \Delta\right) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right]\mathrm{d}t + \sigma S\left(\frac{\partial V}{\partial S} - \Delta\right)\mathrm{d}W^{\mathbb{P}}. \end{aligned}} \tag{3.7}$$

The portfolio, although fully defined in terms of stocks and option, has random fluctuations, governed by Brownian motion $W^{\mathbb{P}}$. By choosing

$$\boxed{\Delta = \frac{\partial V}{\partial S},} \tag{3.8}$$

the infinitesimal change of portfolio $\Pi(t,S)$, in time instance $\mathrm{d}t$, is given by:

$$\mathrm{d}\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t, \tag{3.9}$$

which is now deterministic,[1] *as the* $\mathrm{d}W^{\mathbb{P}}$*-terms cancel out*. Moreover, with Δ as defined by (3.8), the resulting dynamics of the portfolio do not contain the drift parameter μ, which drives the stock $S(t)$ under the real-world measure $\mathbb{P}$ in (2.1), anymore. The value of the portfolio still depends on volatility parameter σ, which is the representation of the uncertainty about the future behaviour of the stock prices.

The value of this portfolio should, on average, grow at the same *speed* (i.e., generate the same return) as money placed on a risk-free money-savings account. The bank account $M(t) = M(t_0)\mathrm{e}^{r(t-t_0)}$, is modeled by means of $\mathrm{d}M = rM\mathrm{d}t$, which, for an amount $\Pi \equiv \Pi(t, S)$ can be expressed as,

$$\mathrm{d}\Pi = r\Pi\mathrm{d}t.$$

Here, r corresponds to the constant interest rate on a money-savings account. With the definitions in Equation (3.6) and the definition of Δ in Equation (3.8), the change in portfolio value is written as:

$$\mathrm{d}\Pi = r\left(V - S\frac{\partial V}{\partial S}\right)\mathrm{d}t. \tag{3.10}$$

Equating Equations (3.10) and (3.9), and dividing both sides by $\mathrm{d}t$, gives us the Black-Scholes partial differential equation for the value of the option $V(t, S)$:

$$\boxed{\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - rV = 0.} \tag{3.11}$$

The Black-Scholes equation (3.11) is a parabolic PDE. With the "+"-sign in front of the diffusion term, i.e. $\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}$, the parabolic PDE problem is *well-posed* when it is accompanied by a *final condition*. The natural condition for the Black-Scholes PDE is indeed a final condition, i.e. we know that,

$$V(T, S) = H(T, S),$$

where $H(T, S)$ is the payoff function, as in (3.1) for a call option, or (3.2) for a put option.

Notice that, except for this final condition, so far we did not specify the type of option in the derivation of the Black-Scholes equation. The equation holds for both calls, puts, and also even for some other option types.

PDE (3.11) is defined on a semi-infinite half space $(t, S) = [t_0 = 0, \ldots, T) \cup [0, S \to \infty)$, and there are no natural conditions at the outer boundaries. When we solve the PDE by means of a numerical method, we need to restrict our computations to a finite computational domain, and we thus need to define a finite domain $(t, S) = [t_0 = 0, \ldots, T) \cup [0, S_{max}]$. We then prescribe appropriate boundary conditions at $S = S_{max}$, as well as at $S = 0$, which is easily possible from an economic viewpoint (as we have information about the value of a call or put, when $S(t) = 0$ or when $S(t)$ is large).

[1] By "deterministic" we mean that the value is defined in terms of other variables, that nota bene, are stochastic, but without additional sources of randomness.

By solving the Black-Scholes equation, we thus know the fair value for the option price at any time $t \in [0, T]$, and at any future stock price $S(t) \in [0, S_{max}]$. The solution of the Black-Scholes PDE will be discussed in detail when the Feynman-Kac Theorem is introduced.

Hedge parameters

Next to these option values, other important information for an option writer is found in the so-called *hedge parameters* or the option Greeks. These are the *sensitivities of the option value* with respect to small changes in the problem parameters or in the independent variables, like the asset S, or the volatility σ.

We have already encountered the option delta, $\Delta = \partial V/\partial S$, in (3.8), representing the rate of change of the value of the option with respect to changes in S. In a replicating portfolio, with stocks to cover the option, Δ gives the number of shares that should be kept by the writer for each option issued, in order to cope with the possible exercise by the holder of the contract at time $t = T$. A negative number implies that short-selling of stocks should take place.

The *sensitivity of the option delta* is called the option *gamma*,

$$\boxed{\Gamma := \frac{\partial^2 V}{\partial S^2} = \frac{\partial \Delta}{\partial S}.}$$

A change in delta is an indication for the stability of a hedging portfolio. For small values of gamma, a writer may not yet need to update the portfolio frequently as the impact of changing the number of stocks (Δ) appears small. When the option gamma is large, however, the hedging portfolio appears only free of risk at a short time scale. There are several other important hedging parameters that we will encounter later in this book.

Dividends

The assumption that dividend payment on the stocks does not take place in the derivation of the Black-Scholes equation can be relaxed. As explained, at the time a dividend is paid, there will be a drop in the value of the stock, see also Equation (2.21). The value of an option on a dividend-paying asset is then also affected by these dividend payments, so that the Black-Scholes analysis must be modified to take dividend payments into account. The constant proportional dividend yield of size $qS(t)$ with $q < 1$, as in (2.21), is considered a satisfactory model for options on stock indices that are of European style. The dividend payment also has its effect on the hedging portfolio. Since we receive $qS(t)\mathrm{d}t$ for every asset held and we hold $-\Delta$ of the underlying, the portfolio changes by an amount $-qS(t)\Delta\mathrm{d}t$, i.e. we have to replace (3.7), by

$$\mathrm{d}\Pi = \mathrm{d}V - \Delta\mathrm{d}S - qS\Delta\mathrm{d}t.$$

Based on the same arguments as earlier, we can then derive the Black-Scholes PDE, where a *continuous stream of dividend payments* is modeled, as

$$\boxed{\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}+(r-q)S\frac{\partial V}{\partial S}-rV=0.} \tag{3.12}$$

3.1.3 Martingale approach and option pricing

Based on the properties of martingales, we give an alternative derivation of the Black-Scholes option pricing PDE.

The following pricing problem, under the risk-neutral GBM model defined by (2.1), is connected to satisfying the martingale property, for a discounted option price:

$$\frac{V(t_0,S)}{M(t_0)}=\mathbb{E}^{\mathbb{Q}}\left[\frac{V(T,S)}{M(T)}\Big|\mathcal{F}(t_0)\right], \tag{3.13}$$

with $M(t_0)$ the money-savings account at time t_0, where $M(t_0)=1$, and $\mathcal{F}(t_0)=\sigma(S(s);s\leq t_0)$.

Since financial options are traded products, Equation (3.13) defines a *discounted option contract* to be a martingale, where the typical martingale properties should then be satisfied. These are properties like the expected value of a martingale should be equal to the present value, as this leads to no-arbitrage values when pricing these financial contracts.

We assume the existence of a differentiable function, $\frac{V(t,S)}{M(t)}$, so that:

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{V(T,S)}{M(T)}\Big|\mathcal{F}(t)\right]=\frac{V(t,S)}{M(t)}. \tag{3.14}$$

Using $M\equiv M(t)$ and $V\equiv V(t,S)$, the discounted option value V/M should be a martingale and its dynamics can be found, as follows,

$$\mathrm{d}\left(\frac{V}{M}\right)=\frac{1}{M}\mathrm{d}V-\frac{V}{M^2}\mathrm{d}M=\frac{1}{M}\mathrm{d}V-r\frac{V}{M}\mathrm{d}t. \tag{3.15}$$

Higher-order terms, like $(\mathrm{d}M)^2=O(\mathrm{d}t^2)$, are omitted in (3.15), as they converge to zero rapidly with infinitesimally small time steps.

For the infinitesimal change $\mathrm{d}V$ of $V(t,S)$, we find under measure $\mathbb{Q}$ using Itô's lemma (3.5),

$$\mathrm{d}V=\left(\frac{\partial V}{\partial t}+rS\frac{\partial V}{\partial S}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t+\sigma S\frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{Q}}. \tag{3.16}$$

As $\frac{V(t,S)}{M(t)}$ should be a martingale, Theorem 1.3.5 states that the dynamics of $\frac{V(t,S)}{M(t)}$ *should not contain any* $\mathrm{d}t$*-terms*. Substituting Equation (3.16) into (3.15), setting the term in front of dt equal to zero, yields that in this case,

$$\frac{1}{M}\left(\frac{\partial V}{\partial t}+rS\frac{\partial V}{\partial S}+\frac{1}{2}\sigma^2S^2\frac{\partial^2 V}{\partial S^2}\right)-r\frac{V}{M}=0, \tag{3.17}$$

assuring that dt-terms are zero. Multiplying both sides of (3.17) by M, gives us the Black-Scholes pricing PDE (3.11), on the basis of the assumption that the martingale property should hold. More information on martingale methods can be found in [Pascucci, 2011].

3.2 The Feynman-Kac theorem and the Black-Scholes model

A number of different solution methods for solving the Black-Scholes PDE (3.11) are available. We are particularly interested in the solution via the Feynman-Kac formula, as given in the theorem below, which forms the basis for a closed-form expression for the option value. It can be generalized for pricing derivatives under other asset price dynamics as well, and forms the basis for option pricing by Fourier methods (in Chapter 6) as well as for pricing by means of Monte Carlo methods (in Chapter 9).

There are different versions of the Feynman-Kac theorem, as the theorem has originally been developed in the context of physics applications. We start with a version of the theorem which is related to option pricing for the case discussed so far, in which we deal with a deterministic interest rate r.

Theorem 3.2.1 (Feynman-Kac theorem) *Given the money-savings account, modeled by* $\mathrm{d}M(t)=rM(t)\mathrm{d}t$*, with constant interest rate* r*, let* $V(t,S)$ *be a sufficiently differentiable function of time* t *and stock price* $S=S(t)$*. Suppose that* $V(t,S)$ *satisfies the following partial differential equation, with general drift term,* $\bar{\mu}(t,S)$*, and volatility term,* $\bar{\sigma}(t,S)$*:*

$$\frac{\partial V}{\partial t}+\bar{\mu}(t,S)\frac{\partial V}{\partial S}+\frac{1}{2}\bar{\sigma}^2(t,S)\frac{\partial^2 V}{\partial S^2}-rV=0, \tag{3.18}$$

with a final condition given by $V(T,S)=H(T,S)$*. The solution* $V(t,S)$ *at any time* $t<T$ *is then given by:*

$$V(t,S)=\mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[H(T,S)\big|\mathcal{F}(t)\right]=:M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{H(T,S)}{M(T)}\big|\mathcal{F}(t)\right]$$

where the expectation is taken under the measure $\mathbb{Q}$*, with respect to a process* S*, which is defined by:*

$$\mathrm{d}S(t)=\bar{\mu}(t,S)\mathrm{d}t+\bar{\sigma}(t,S)\mathrm{d}W^{\mathbb{Q}}(t),\quad t>t_0. \tag{3.19}$$

Proof We present a short outline of the proof. Consider the term

$$\frac{V(t,S)}{M(t)}=\mathrm{e}^{-r(t-t_0)}V(t,S),$$

for which we find the dynamics:

$$\mathrm{d}(\mathrm{e}^{-r(t-t_0)}V(t,S)) = V(t,S)\mathrm{d}(\mathrm{e}^{-r(t-t_0)}) + \mathrm{e}^{-r(t-t_0)}\mathrm{d}V(t,S). \tag{3.20}$$

Using $V := V(t,S)$, $S := S(t)$, $\bar{\mu} := \bar{\mu}(t,S)$, $\bar{\sigma} := \bar{\sigma}(t,S)$ and $W^{\mathbb{Q}} := W^{\mathbb{Q}}(t)$, and applying Itô's lemma to $V := V(t,S)$, we obtain:

$$\mathrm{d}V = \left(\frac{\partial V}{\partial t} + \bar{\mu}\frac{\partial V}{\partial S} + \frac{1}{2}\bar{\sigma}^2\frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t + \bar{\sigma}\frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{Q}}.$$

By multiplying (3.20) by $\mathrm{e}^{r(t-t_0)}$ and substituting the above expression in it, we find:

$$\mathrm{e}^{r(t-t_0)}\mathrm{d}(\mathrm{e}^{-r(t-t_0)}V) = \underbrace{\left(\frac{\partial V}{\partial t} + \bar{\mu}\frac{\partial V}{\partial S} + \frac{1}{2}\bar{\sigma}^2\frac{\partial^2 V}{\partial S^2} - rV\right)}_{=0}\mathrm{d}t + \bar{\sigma}\frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{Q}},$$

in which the $\mathrm{d}t$-term is equal to zero, because of the original PDE (3.18) which equals zero in the theorem. Integrating both sides gives us,

$$\begin{aligned}\int_{t_0}^{T}\mathrm{d}(\mathrm{e}^{-r(t-t_0)}V(t,S)) &= \int_{t_0}^{T}\mathrm{e}^{-r(t-t_0)}\bar{\sigma}\frac{\partial V}{\partial S}\,\mathrm{d}W^{\mathbb{Q}}(t) \quad \Leftrightarrow \\ \mathrm{e}^{-r(T-t_0)}V(T,S) - V(t_0,S) &= \int_{t_0}^{T}\mathrm{e}^{-r(t-t_0)}\bar{\sigma}\frac{\partial V}{\partial S}\,\mathrm{d}W^{\mathbb{Q}}(t).\end{aligned}$$

We now take the expectation at both sides of this equation, with respect to the $\mathbb{Q}$-measure, and rewrite as follows:

$$V(t_0,S) = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-r(T-t_0)}V(T,S)\Big|\mathcal{F}(t_0)\right] - \mathbb{E}^{\mathbb{Q}}\left[\int_{t_0}^{T}\mathrm{e}^{-r(t-t_0)}\bar{\sigma}\frac{\partial V}{\partial S}\,\mathrm{d}W^{\mathbb{Q}}(t)\Big|\mathcal{F}(t_0)\right].$$

According to the properties of an Itô integral, as defined in (1.23), $I(t_0) = 0$ and $I(t) = \int_0^t g(s)\mathrm{d}W^{\mathbb{Q}}(s)$ is a martingale, so that we have that $\mathbb{E}^{\mathbb{Q}}[I(t)|\mathcal{F}(t_0)] = 0$ for all $t \geq t_0$.

The expectation of the Itô integral which is based on Wiener process $W^{\mathbb{Q}}(t)$ is therefore equal to 0, so it follows that

$$V(t_0,S) = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-r(T-t_0)}V(T,S)\big|\mathcal{F}(t_0)\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-r(T-t_0)}H(T,S)\big|\mathcal{F}(t_0)\right].$$

This concludes this outline of the proof. ■

By the Feynman-Kac theorem we can thus translate the problem of *solving the Black-Scholes PDE*, which is recovered by specific choices for drift term $\bar{\mu}(t,S) = rS$ and diffusion term $\bar{\sigma}(t,S) = \sigma S$, to the calculation of an expectation of a discounted payoff function under the $\mathbb{Q}$-measure.

A comprehensive discussion on the Feynman-Kac theorem and its various properties have been provided, for example, in [Pelsser, 2000].

Remark 3.2.1 (Log coordinates under the GBM model)
The Feynman-Kac theorem, as presented in Theorem 3.2.1, also holds true in logarithmic coordinates, i.e., when $X(t) = \log S(t)$. With this transformation of variables, the resulting log-transformed Black-Scholes PDE reads

$$\frac{\partial V}{\partial t} + r\frac{\partial V}{\partial X} + \frac{1}{2}\sigma^2\left(-\frac{\partial V}{\partial X} + \frac{\partial^2 V}{\partial X^2}\right) - rV = 0, \tag{3.21}$$

which comes with the transformed pay-off function, $V(T, X) = H(T, X)$. The Feynman-Kac theorem gives us, for the specific choice of $\bar{\mu} = (r - \frac{1}{2}\sigma^2)$ and $\bar{\sigma} = \sigma$, the following representation of the option value,

$$V(t, X) = \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[H(T, X)\middle|\mathcal{F}(t)\right],$$

where $H(T, X)$ represents the terminal condition in log-coordinates $X(t) = \log S(t)$, and

$$\mathrm{d}X(t) = \left(r - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}W^{\mathbb{Q}}(t). \tag{3.22}$$

▲

Note that we have used essentially the same *notation*, $V(\cdot,\cdot)$, for the option value when the asset price is an independent variable, i.e. $V(t, S)$, as when using the log-asset price as the independent variable, i.e. $V(t, X)$. From the context it will become clear how the option is defined.

Regarding the terminal condition, we use $H(\cdot,\cdot)$ also for the call and put payoff functions in log-coordinates, i.e. $H(T, X) = H_c(T, X) = \max(\mathrm{e}^X - K, 0)$, $H(T, X) = H_p(T, X) = \max(K - \mathrm{e}^X, 0)$, respectively.

Throughout the book, we will often use $\tau = T - t$, which is common practice, related to the fact that people are simply used to "time running forward" while we know the payoff function at $\tau = 0$ we do not know the value at $\tau = T - t_0$. Of course, with $\mathrm{d}\tau = -\mathrm{d}t$, a minus sign is added to a derivative with respect to τ.

3.2.1 Closed-form option prices

Some results will follow, where it is shown that for some payoff functions under the Black-Scholes dynamics an analytic solution, which is obtained via the Feynman-Kac theorem, is available for European options.

Theorem 3.2.2 (European call and put option) *A closed-form solution of the Black-Scholes PDE (3.11) for a European call option with a constant strike price K can be derived, with $H_c(T, S) = \max(S(T) - K, 0)$. The option value at any time t (in particular, also at time now, i.e. $t = t_0$), under the Black-Scholes model, of a call option can be written as:*

$$\begin{aligned} V_c(t, S) &= \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\max(S(T) - K, 0)\middle|\mathcal{F}(t)\right] \\ &= \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[S(T)\mathbb{1}_{S(T)>K}\middle|\mathcal{F}(t)\right] - \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[K\mathbb{1}_{S(T)>K}\middle|\mathcal{F}(t)\right], \end{aligned}$$

The solution is then given by:

$$V_c(t,S) = S(t)F_{\mathcal{N}(0,1)}(d_1) - K\mathrm{e}^{-r(T-t)}F_{\mathcal{N}(0,1)}(d_2), \tag{3.23}$$

with

$$\begin{aligned} d_1 &= \frac{\log\frac{S(t)}{K} + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \\ d_2 &= \frac{\log\frac{S(t)}{K} + (r-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}, \end{aligned}$$

and $F_{\mathcal{N}(0,1)}(\cdot)$ the cumulative distribution function of a standard normal variable.
It is easy to show that this is indeed the solution by substitution of the solution is the Black-Scholes PDE.
Similarly, for a European put option, we find:

$$V_p(t,S) = K\mathrm{e}^{-r(T-t)}F_{\mathcal{N}(0,1)}(-d_2) - S(t)F_{\mathcal{N}(0,1)}(-d_1),$$

with d_1 and d_2 as defined above. This solution can directly be found by writing out the integral form in (3.22), based on the log-transformed asset process.

Definition 3.2.1 (Closed-form expression for Δ) *Greek delta is given by $\Delta := \frac{\partial V}{\partial S}$, so by differentiation of the option price, provided in Theorem 3.2.2, we obtain,*

- *For call options,*

$$\Delta = \frac{\partial}{\partial S}V_c(t,S) = F_{\mathcal{N}(0,1)}(d_1). \tag{3.24}$$

  with d_1 as in Theorem 3.2.2.

- *For put options,*

$$\begin{aligned} \Delta &= \frac{\partial}{\partial S}V_p(t,S) \\ &= -F_{\mathcal{N}(0,1)}(-d_1) = F_{\mathcal{N}(0,1)}(d_1) - 1, \end{aligned} \tag{3.25}$$

In particular, we are interested in the values of the Greeks, at $t = t_0$, i.e. for $S = S_0$. ◀

Example 3.2.1 (Black-Scholes solution) In this example, we will present a typical surface of option values. We will use a call option for this purpose, and the following parameter set:

$$S_0 = 10, r = 0.05, \sigma = 0.4, T = 1, K = 10.$$

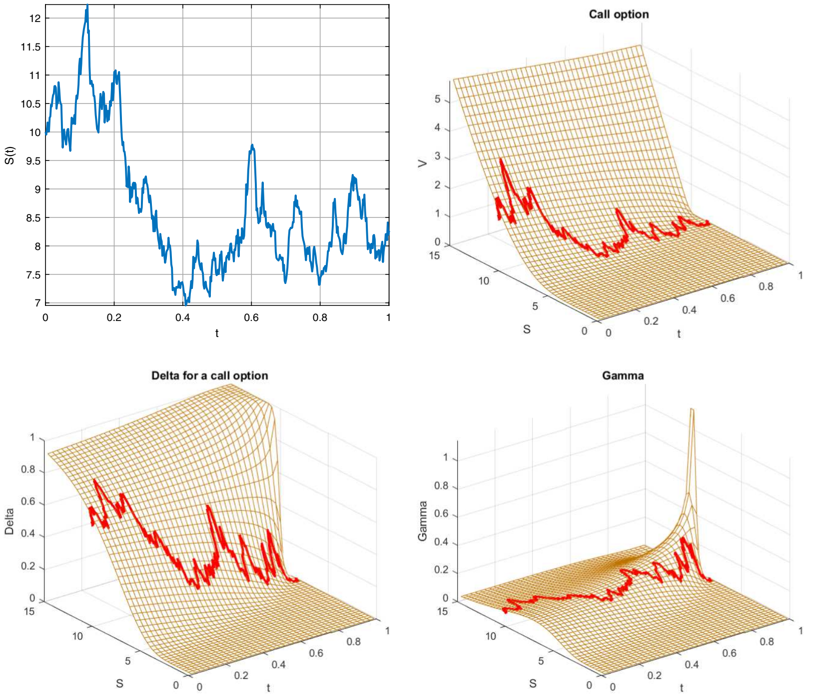


Figure 3.3: *Example of a call option with exercise date $T = 1$, the option surface $V_c(t, S)$ in the (t, S)-domain is computed. An asset path intersects the surface, so that at each point (t, S), the corresponding option value $V_c(t, S)$ can be read. The surfaces for the hedge parameters delta and gamma are also presented, with the projected stock price.*

Once the problem parameters have been defined, we can calculate the option values at any time t and stock price $S(t)$, by the solution in (3.23). Visualizing these call option values $V_c(t, S)$ in an (t, S)-plane, gives us the surface in Figure 3.3. From the closed-form solution, calculated by the Black-Scholes formula, we can also easily determine the first and second derivatives with respect to the option value, i.e., the option delta and gamma. They are presented in the same figure. When the stock price then moves at some time, $t > t_0$, we can "read" the corresponding option values, and also the option Greeks, see also Figure 3.3.

First of all, the option value surface is computed, and then the observed stock prices are projected on top of this surface. We also display the corresponding put option surface $V_p(t, S)$ in the (t, S)-plane in Figure 3.4.

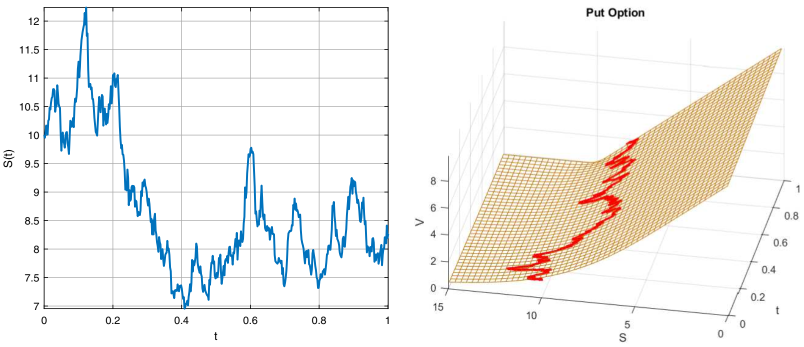


Figure 3.4: *Example of a put option, the option surface $V_p(t, S)$ in the (t, S)-domain is displayed with an asset path that intersects the surface.*

Digital option

As a first option contract away from the standard European put and call options, we discuss digital options, that are also called cash-or-nothing options. Digital options are popular for hedging and speculation. They are useful as building blocks for more complex option contracts. Consider the payoff of a cash-or-nothing digital call option, whose value is equal to 0 for $S(T) \leq K$ and its value equals K when $S(T) > K$, see Figure 3.5 (left side). The price at time t is then given by:

$$\begin{aligned} V(t, S) &= \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[K\mathbb{1}_{S(T)>K}\middle|\mathcal{F}(t)\right] \\ &= \mathrm{e}^{-r(T-t)}K\mathbb{Q}\left[S(T) > K\right], \end{aligned}$$

as $\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{S(T)>K}\right] := \mathbb{Q}\left[S(T) > K\right]$. Hence, the value of a cash-or-nothing call under the Black-Scholes dynamics, is given by

$$V(t, S) = K\mathrm{e}^{-r(T-t)}F_{\mathcal{N}(0,1)}(d_2), \tag{3.26}$$

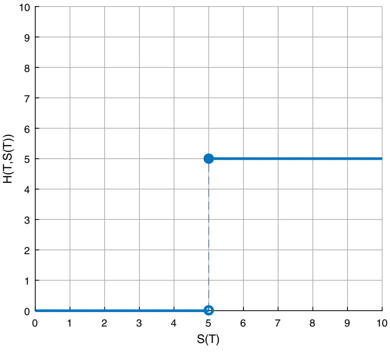
with d_2 defined in Theorem 3.2.2. Similarly, we can derive an expression for the solution of an *asset-or-nothing* option with payoff function in Figure 3.5 (right side).

3.2.2 Green's functions and characteristic functions

Based on the Feynman-Kac theorem we obtain interesting results and solutions by a clever choice of final condition $H(\cdot, \cdot)$ [Heston, 1993]. If, for example,[2]

$$H(T, X) := \mathbb{1}_{X(T)\geq \log K},$$

[2] Of course, it holds that $\mathbb{1}_{X(T)\geq \log K} = \mathbb{1}_{S(T)\geq K}$.



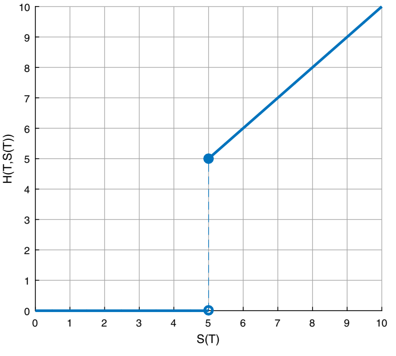


Figure 3.5: *Payoff functions for the cash-or-nothing (left), with* $K\mathbb{1}_{S(T)\geq K}$ *and* $K = 5$, *and asset-or-nothing (right), with* $S(T)\mathbb{1}_{S(T)\geq K}$ *and* $K = 5$, *call options.*

is chosen, the solution of the corresponding PDE is the *conditional probability* that $X(T)$ is greater than $\log K$.

As another example, we consider the following form of the final condition[a]:

$$\phi_X(u;T,T) = H(T,X) = \mathrm{e}^{iuX(T)}.$$

By the Feynman-Kac theorem, the solution for $\phi_X := \phi_X(u;t,T)$ is then given by:

$$\phi_X(u;t,T) = \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{iuX(T)}\middle|\mathcal{F}(t)\right], \tag{3.27}$$

which is by definition *the discounted characteristic function of* $X(t)$. We thus also know that $\phi_X(u;t,T)$ is a solution of the corresponding PDE:

$$\frac{\partial\phi_X}{\partial t} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial\phi_X}{\partial X} + \frac{1}{2}\sigma^2\frac{\partial^2\phi_X}{\partial X^2} - r\phi_X = 0, \tag{3.28}$$

with terminal condition $\phi_X(u;T,T) = \mathrm{e}^{iuX(T)}$.

[a]In order to emphasize that we will deal here with a characteristic function we use $\phi_X(u;t,T)$ instead of H.

In Chapter 6, we will work with the discounted characteristic function under the log-transformed GBM asset process, which is given by

$$\phi_X(u;t,T) = \exp\left[\left(r - \frac{\sigma^2}{2}\right)iu\tau - \frac{1}{2}\sigma^2u^2\tau - r\tau + iuX\right], \tag{3.29}$$

with $\tau = T - t$. For a stochastic process $X(t)$, $t > 0$, we denote its characteristic function by $\phi_X(u;t,T)$ or simply by $\phi_X(u,X,\tau)$ with $\tau = T - t$.

For a large number of stochastic underlying asset models for which the density is not known, the *characteristic function* (see the definition in Equation (1.4)), or the discounted characteristic function, is known, and this function also uniquely determines the transition density function.

Green's function

With the Feynman-Kac Theorem, we are able to make connections between a PDE and the integral formulation of a solution. By this, we can relate to well-known PDE terminology, for example, we can determine the Green's function which is related to the integral formulation of the solution of the PDE. We will see that it is, in fact, related to the asset price's density function. Moreover, with the Feynman-Kac Theorem we can derive the characteristic function, which we will need for option pricing with Fourier techniques in the chapters to come.

Based on the Feynman-Kac theorem, under the assumption of a constant interest rate, we can write the value of a European option as,

$$V(t_0, x) = \mathrm{e}^{-r(T-t_0)} \int_{\mathbb{R}} H(T, y) f_X(T, y; t_0, x) \mathrm{d}y, \tag{3.30}$$

where $f_X(T, y; t_0, x)$ represents the *transition probability density function* in log-coordinates from state $\log S(t_0) = X(t_0) = x$ at time t_0, to state $\log S(T) = X(T) =: y$ at time T.

De facto, if we take $t_0 = 0$ and initial value x constant, we have here a marginal probability density function.

Taking a closer look at Equation (3.30), we recognize for the PDE (3.21) the *fundamental solution*, which is denoted here by $G_X(T, y; t_0, x) := \mathrm{e}^{-r(T-t_0)} f_X(T, y; t_0, x)$, times the final condition $V(T, y) \equiv H(T, y)$, in the form of a convolution. The solution to the PDE can thus be found based on this fundamental solution by means of the Feynman-Kac theorem. The fundamental solution $\mathrm{e}^{-r(T-t_0)} f_X(T, y; t_0, x)$ can be interpreted as the *Green's function*, which is connected to the following parabolic final value problem:

$$-\frac{\partial f_X}{\partial \tau} + r\frac{\partial f_X}{\partial x} + \frac{1}{2}\sigma^2 \left(-\frac{\partial f_X}{\partial x} + \frac{\partial^2 f_X}{\partial x^2}\right) - r f_X = 0, \tag{3.31}$$
$$f_X(T, y; T, y) = \delta(y = \log K),$$

with $\delta(y = \log K)$ the Dirac delta function, which is nonzero for $y := X(T) = \log K$ and is 0 otherwise, and its integral equals one, see (1.17).

The Green's function, related to the log-space Black-Scholes PDE, is equal to a *discounted* normal probability density function, i.e.,

$$G_X(T, X; t_0, x) = \frac{\mathrm{e}^{-r(T-t_0)}}{\sigma\sqrt{2\pi(T-t_0)}} \exp\left(-\frac{(X - x - (\mu - \frac{1}{2}\sigma^2)(T-t_0))^2}{2\sigma^2(T-t_0)}\right). \tag{3.32}$$

The Green's function in derivative pricing, i.e. the discounted risk neutral probability density, has received the name *Arrow-Debreu security* in finance. We can write out the Green's function for the Black-Scholes operator *in the original* (S,t) *coordinates* explicitly, i.e. for $y = S(T), x = S(t_0)$, we find:

$$G_S(T,y;t_0,x) = \frac{\mathrm{e}^{-r(T-t_0)}}{\sigma y\sqrt{2\pi(T-t_0)}} \exp\left(-\frac{\left(\log\left(\frac{y}{x}\right) - (r - \frac{1}{2}\sigma^2)(T-t_0)\right)^2}{2\sigma^2(T-t_0)}\right), \tag{3.33}$$

see also Equation (2.20). Density functions and Green's functions are known in closed-form only for certain stochastic models under basic underlying dynamics.

Arrow-Debreu security and the market implied density

A *butterfly spread option* is an option strategy which may give the holder a predetermined profit when the stock price movement stays within a certain range of stock prices. With $S(t_0) = K$, the butterfly spread option is defined by two long call options, one call with strike price $K_1 = K - \Delta K$, another call with strike price $K_3 = K + \Delta K$, $\Delta K > 0$, and by simultaneously two short call options with strike price $K_2 = K$. All individual options have the same maturity time $t = T$, with payoff,

$$V_B(T,S;K_2,T) = V_c(T,S;K_1,T) + V_c(T,S;K_3,T) - 2V_c(T,S;K_2,T), \tag{3.34}$$

see Figure 3.6. At the expiry date $t = T$, this construction will be profitable if the asset $S(T) \in [K - \Delta K, K + \Delta K]$; the butterfly spread has a zero value if $S(T) < K - \Delta K$ or $S(T) > K + \Delta K$. Since the butterfly option is a linear combination of individual option payoffs, it satisfies the Black-Scholes PDE, when we assume that stock $S(t)$ is modeled by a GBM.[3]

The butterfly spread option at time $t = t_0$ can be calculated as[4]

$$V_B(t_0,S;K,T) = \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\big[V_B(T,S;K,T)|\mathcal{F}(t_0)\big], \tag{3.35}$$

based on the Feynman-Kac theorem.

This butterfly option is related to the so-called *Arrow-Debreu security* $V_{AD}(t,S)$, which is an imaginary security, that only pays out one unit at maturity date $t = T$, if $S(T) \equiv K$. The payoff is zero otherwise, i.e. $V_{AD}(T,S) = \delta(S(T) - K)$.

---

[3]Actually, the result here holds for any dynamics of S, it is not specific to GBM.

[4]We explicitly add the dependency of the option on the strike price K in the option argument list.

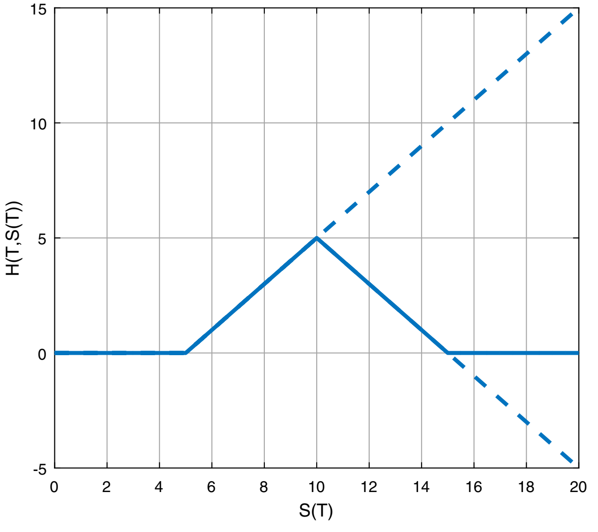


Figure 3.6: *Butterfly spread payoff function, with $K_1 = 5$, $K_2 = 10$ and $K_3 = 15$.*

The Arrow-Debreu security can be constructed by means of three European calls, like the butterfly spread option V_B:

$$
\begin{aligned}
V_{AD}(T, S; K, T) := \frac{1}{\Delta K^2}&[V_c(T, S; K - \Delta K, T) + V_c(T, S; K + \Delta K, T)\\
&- 2V_c(T, S; K, T)],
\end{aligned}
$$

for small $\Delta K > 0$, with $V_c(t, S; K, T)$ representing the value of a European call at time t, with strike K, expiry date T.

Note that for $\Delta K \to 0$, we have

$$
V_{AD}(T, S; K, T) = \frac{\partial^2 V_c(T, S; K, T)}{\partial K^2}.
$$

On the other hand, in the limit $\Delta K \to 0$, we also find

$$
\begin{aligned}
\lim_{\Delta K \to 0} V_{AD}(t_0, S; K, T) &= \lim_{\Delta K \to 0} \mathrm{e}^{-r(T-t_0)} \mathbb{E}^{\mathbb{Q}}\big[V_{AD}(T, S; K, T) | \mathcal{F}(t_0)\big]\\
&= \mathrm{e}^{-r(T-t_0)} \mathbb{E}^{\mathbb{Q}}\big[\delta(S(T) - K) | \mathcal{F}(t_0)\big]\\
&= \mathrm{e}^{-r(T-t_0)} \int_0^\infty f_S(T, S(T); t_0, S_0) \cdot \delta(S(T) - K) \mathrm{d}S(T)\\
&= \mathrm{e}^{-r(T-t_0)} f_S(T, K; t_0, S_0), \qquad (3.36)
\end{aligned}
$$

where $\delta(S(T) - K)$ is the Dirac delta function which is only nonzero when $S(T) = K$, $f_S(T, K; t_0, S_0)$ is the *transition risk-neutral density* with $S(T) \equiv K$. Transition densities model the evolution of a probability density through time, for example, between the time points s and t. Here we set time point $s = t_0$, which actually gives rise to the marginal distribution (the density of the marginal distribution is a special marginal density, for $s = t_0$).

The above expression shows a direct relation between the density and the butterfly spread option. With this connection, we can examine whether well-known properties of density functions (nonnegativity, integrating to one) are satisfied, for example for option prices observed in the financial markets. This property will be examined in Chapter 4 where local volatility asset models will be introduced.

3.2.3 Volatility variations

A generalization of the Black-Scholes model, which was already introduced in Equation (2.23) was to prescribe a time-dependent volatility coefficient $\sigma(t)$, instead of constant volatility.

Log-transformed GBM asset prices are normally distributed and in that case only the first two moments have to match to guarantee equality in distribution. By means of matching the moments of the GBM process and of a GBM with time-dependent volatility, it was shown in Example 2.2.1 under which conditions a model with time-dependent volatility and one with a constant volatility have identical first two moments. This implies, in fact, that European option values under a Black-Scholes model with a time-dependent volatility function are identical to the corresponding option values that are obtained by a log-transformed model based on the following *time-averaged, constant volatility function*,

$$\sigma_* = \sqrt{\frac{1}{T}\int_0^T \sigma^2(t)\mathrm{d}t}.$$

The valuation of European options under the GBM asset model with a time-dependent volatility can thus be performed also by using a constant volatility coefficient σ_*, as derived in (2.25), because the first two moments are identical. The difference between the two models is subtle and can only be observed when considering pricing exotic options (e.g. for path-dependent options).

Iterated expectations and stochastic volatility

We present another basic and often used application of the tower property of the expectation in finance. In this example we assume the following SDE for a stock price,

$$\mathrm{d}S(t) = rS(t)\mathrm{d}t + Y(t)S(t)\mathrm{d}W^{\mathbb{Q}}(t),$$

where $Y(t)$ represents a certain *stochastic volatility process* which has, for example, a lognormal distribution. After standard calculations, we obtain the following solution for $S(T)$, given by:

$$S(T) = S_0 \exp\left(\int_{t_0}^{T}\left(r - \frac{1}{2}Y^2(t)\right)\mathrm{d}t + \int_{t_0}^{T} Y(t)\mathrm{d}W^{\mathbb{Q}}(t)\right). \tag{3.37}$$

As Expression (3.37) contains integrals of the process $Y(t)$, it is nontrivial to determine a closed-form solution for the value of a European option.[5] A possible solution for the pricing problem is to use the *tower property of iterated expectations*, to determine the European option prices, conditioned on "realizations" of the volatility process $Y(t)$.

By the tower property of expectations, using $\mathbb{E} = \mathbb{E}^{\mathbb{Q}}$, the European call value can be reformulated as a discounted expectation with[6]:

$$\begin{aligned}&\mathbb{E}\left[\max\left(S(T) - K, 0\right)\Big|\mathcal{F}(t_0)\right]\\ &\quad= \mathbb{E}\left[\mathbb{E}\left[\max\left(S(T) - K, 0\right)\Big|Y(t), t_0 \le t \le T\right]\Big|\mathcal{F}(t_0)\right].\end{aligned} \tag{3.38}$$

Conditioned on the realizations of the variance process, the calculation of the inner expectation is equivalent to the Black-Scholes solution with a time-dependent volatility, i.e. for given realizations of $Y(t)$, $t_0 \le t \le T$, the asset value $S(T)$ in (3.37) is given by:

$$S(T) = S(t_0)\exp\left(\left(r - \frac{1}{2}\sigma_*^2\right)(T - t_0) + \sigma_*(W^{\mathbb{Q}}(T) - W^{\mathbb{Q}}(t_0))\right),$$

with

$$\sigma_*^2 = \frac{1}{T - t_0}\int_{t_0}^{T} Y^2(t)\mathrm{d}t.$$

The solution of the inner expectation is then given by:

$$\mathbb{E}\left[\max\left(S(T) - K, 0\right)\Big|\{Y(t)\}_{t_0}^{T}\right] = S(t_0)\mathrm{e}^{r(T-t_0)}F_{\mathcal{N}(0,1)}(d_1) - KF_{\mathcal{N}(0,1)}(d_2),$$

with

$$d_1 = \frac{\log\frac{S(t_0)}{K} + (r + \frac{1}{2}\sigma_*^2)(T - t_0)}{\sigma_*\sqrt{T - t_0}}, \quad d_2 = d_1 - \sigma_*\sqrt{T - t_0},$$

$F_{\mathcal{N}(0,1)}$ being the standard normal cumulative distribution function.

We can substitute these results into Equation (3.38), giving:

$$\begin{aligned}\mathbb{E}\left[\max\left(S(T) - K, 0\right)\right] &= \mathbb{E}\left[S(t_0)\mathrm{e}^{r(T-t_0)}F_{\mathcal{N}(0,1)}(d_1) - KF_{\mathcal{N}(0,1)}(d_2)\right]\\ &= S(t_0)\mathrm{e}^{r(T-t_0)}\mathbb{E}\left[F_{\mathcal{N}(0,1)}(d_1)\right] - K\mathbb{E}\left[F_{\mathcal{N}(0,1)}(d_2)\right].\end{aligned}$$

[5] If $Y(t)$ is normally distributed, the model is called the Schöbel-Zhu model, and if $Y(t)$ follows a Cox-Ingersol-Ross (CIR) square-root process, the system is called the Heston model. They are introduced in Chapter 8.

[6] The discount term $M(T)$ is omitted, only the expectation is displayed to save some space. The interest rates in this model are constant and do not influence the final result.

The option pricing problem under these nontrivial asset dynamics has been transformed into the calculation of an expectation of a normal CDF. The difficult part of this expectation is that both arguments of the CDF, d_1 and d_2, are functions of σ_*, which itself is a function of process $Y(t)$. One possibility to deal with the expectation is to use Monte-Carlo simulation, to be discussed in Chapter 9, another is to introduce an approximation, for example, $F_{\mathcal{N}(0,1)}(x) \approx g(x)$ (as proposed in [Piterbarg, 2005a], where $g(x)$ is a function, like $g(x) = a + b\mathrm{e}^{-cx}$).

3.3 Delta hedging under the Black-Scholes model

In this section we discuss some insightful details of the delta hedging strategy and of re-balancing a financial portfolio. The main concept of hedging is to eliminate risk or at least to reduce it when complete elimination is not possible.

The easiest way to eliminate the risk is to offset a trade. A so-called *back-to-back transaction* eliminates the risk associated with market movements. The main idea of such a transaction is as follows. Suppose a financial institution sells some financial derivative to a counterparty. The value of this derivative will vary, depending on market movements. One way to eliminate such movements is to buy "exactly the same", or a very similar, derivative from another counterparty. One may wonder about buying and selling the same derivatives at the same time, but a profit may be made from selling the derivative with some additional premium which is added to the price which is paid.

In Subsection 3.1.2 we discussed the concept of the replicating portfolio, under the Black-Scholes model, where the uncertainty of a European option was eliminated by the Δ hedge.

Based on the same strategy, consider the following portfolio:

$$\Pi(t, S) = V(t, S) - \Delta S. \tag{3.39}$$

The objective of delta hedging is that the value of the portfolio does not change when the underlying asset moves, so the derivative of portfolio $\Pi(t, S)$ w.r.t S needs to be equal to 0, i.e.,

$$\frac{\partial \Pi(t, S)}{\partial S} = \frac{\partial V(t, S)}{\partial S} - \Delta = 0 \Rightarrow \Delta = \frac{\partial V}{\partial S}, \tag{3.40}$$

with $V = V(t, S)$, and Δ given by Equation (3.24).

Suppose we sold a call $V_c(t_0, S)$ at time t_0, with maturity T and strike K. By selling, we obtained a cash amount equal to $V_c(t_0, S)$ and perform a *dynamic* hedging strategy until time T. Initially, at the inception time, we have

$$\Pi(t_0, S) := V_c(t_0, S) - \Delta(t_0)S_0.$$

This value may be negative when $\Delta(t_0)S_0 > V_c(t_0, S)$. If funds are needed for buying $\Delta(t)$ shares, we make use of a *funding account*, $\mathrm{PnL}(t) \equiv \mathrm{P\&L}(t)$. $\mathrm{PnL}(t)$ represents the total value of the option sold and the hedge, and it keeps track of the changes in the asset value $S(t)$.

Typically the funding amount $\Delta(t_0)S_0$ is then obtained from a trading desk of a treasury department.

Every day we may need to re-balance the position and hedge the portfolio. At some time $t_1 > t_0$, we then receive (or pay) interest over the time period $[t_0, t_1]$, which will amount to $\text{P\&L}(t_0)\mathrm{e}^{r(t_1-t_0)}$. At t_1 we have $\Delta(t_0)S(t_1)$ which may be sold, and we will update the hedge portfolio. Particularly, we purchase $\Delta(t_1)$ stocks, costing $-\Delta(t_1)S(t_1)$. The overall $\text{P\&L}(t_1)$ account will become:

$$\text{P\&L}(t_1) = \underbrace{\text{P\&L}(t_0)\mathrm{e}^{r(t_1-t_0)}}_{\text{interest}} \underbrace{-\left(\Delta(t_1) - \Delta(t_0)\right) S(t_1)}_{\text{borrow}}. \tag{3.41}$$

Assuming a time grid with $t_i = i\frac{T}{m}$, the following recursive formula for the m time steps is obtained,

$$\begin{aligned} \text{P\&L}(t_0) &= V_c(t_0, S) - \Delta(t_0)S(t_0), \\ \text{P\&L}(t_i) &= \text{P\&L}(t_{i-1})\mathrm{e}^{r(t_i-t_{i-1})} - \left(\Delta(t_i) - \Delta(t_{i-1})\right) S(t_i), \quad i = 1, \ldots, m-1. \end{aligned} \tag{3.42}$$

At the option maturity time T, the option holder may exercise the option or the option will expire worthless. As the option writer, we will thus encounter a cost equal to the option's payoff at $t_m = T$, i.e. $\max(S(T) - K, 0)$. On the other hand, at maturity time we own $\Delta(t_m)$ stocks, that may be sold in the market. The value of the portfolio at maturity time $t_m = T$ is then given by:

$$\begin{aligned} \text{P\&L}(t_m) = {} & \text{P\&L}(t_{m-1})\mathrm{e}^{r(t_m-t_{m-1})} \\ & - \underbrace{\max(S(t_m) - K, 0)}_{\text{option payoff}} + \underbrace{\Delta(t_{m-1})S(t_m)}_{\text{sell stocks}}. \end{aligned} \tag{3.43}$$

In a perfect world, with continuous re-balancing, the $\text{P\&L}(T)$ would equal zero on average, i.e. $\mathbb{E}[\text{P\&L}(T)] = 0$. One may question the reasoning behind dynamic hedging if the profit made by the option writer on average equals zero. The profit in option trading, especially with OTC transactions, is to charge an additional fee (often called a "*spread*") at the start of the contract. At time t_0 the cost for the client is not $V_c(t_0, S)$ but $V_c(t_0, S) + \textit{spread}$, where $\textit{spread} > 0$ would be the profit for the writer of the option.

Example 3.3.1 (Expected P&L(T)) We will give an example where $\mathbb{E}[\text{P\&L}(T)|\mathcal{F}(t_0)] = 0$. We consider a three-period case, with $t_0, t_1, t_2 := T$, and $\Delta(t_i)$ is a deterministic function, as in the Black-Scholes hedging case. At t_0 an option is sold, with expiry date t_2 and strike price K. In a three-period setting, we have three equations for $\text{P\&L}(t)$ related to the initial hedging, the re-balancing and the final hedging, i.e.,

$$\begin{aligned} \text{P\&L}(t_0) &= V_c(t_0, S) - \Delta(t_0)S(t_0), \\ \text{P\&L}(t_1) &= \text{P\&L}(t_0)\mathrm{e}^{r(t_1-t_0)} - \left(\Delta(t_1) - \Delta(t_0)\right) S(t_1), \\ \text{P\&L}(t_2) &= \text{P\&L}(t_1)\mathrm{e}^{r(t_2-t_1)} - \max(S(t_2) - K, 0) + \Delta(t_1)S(t_2). \end{aligned}$$

After collecting all terms, we find,

$$\begin{aligned}\text{P\&L}(t_2) = \Big[&(V_c(t_0,S) - \Delta(t_0)S(t_0))\,\mathrm{e}^{r(t_2-t_0)}\\ &- (\Delta(t_1) - \Delta(t_0))\,S(t_1)\mathrm{e}^{r(t_2-t_1)}\Big] - \max(S(t_2)-K,0) + \Delta(t_1)S(t_2).\end{aligned}$$

By the definition of a call option, we also have,

$$\mathbb{E}\left[\max(S(t_2)-K,0)\big|\mathcal{F}(t_0)\right] = \mathrm{e}^{r(t_2-t_0)}V_c(t_0,S), \tag{3.44}$$

and because the discounted stock price, under the risk-neutral measure, is a martingale, $\mathbb{E}[S(t)|\mathcal{F}(s)] = S(s)\mathrm{e}^{r(t-s)}$, the expectation of the P&L is given by:

$$\begin{aligned}\mathbb{E}[\text{P\&L}(t_2)|\mathcal{F}(t_0)] = \big(&V_c(t_0,S) - \mathbb{E}[\max(S(t_2)-K,0)|\mathcal{F}(t_0)]\\ &+ \Delta(t_1)\mathbb{E}[S(t_2)|\mathcal{F}(t_0)] - \Delta(t_0)S(t_0)\big)\cdot \mathrm{e}^{r(t_2-t_0)}\\ &- (\Delta(t_1)-\Delta(t_0))\,\mathbb{E}[S(t_1)|\mathcal{F}(t_0)]\mathrm{e}^{r(t_2-t_1)}.\end{aligned}$$

Using the relation

$$\mathbb{E}[S(t_1)|\mathcal{F}(t_0)]\mathrm{e}^{r(t_2-t_1)} = \mathbb{E}[S(t_2)|\mathcal{F}(t_0)] = S(t_0)\mathrm{e}^{r(t_2-t_0)},$$

and by (3.44), the expression simplifies to,

$$\mathbb{E}[\text{P\&L}(t_2)|\mathcal{F}(t_0)] = -\Delta(t_1)S(t_0)\mathrm{e}^{r(t_2-t_0)} + \Delta(t_1)\mathbb{E}[S(t_2)|\mathcal{F}(t_0)] = 0. \qquad \blacklozenge$$

Example 3.3.2 (Dynamic hedge experiment, Black-Scholes model)
In this experiment we perform a dynamic hedge for a call option under the Black-Scholes model. For the asset price, the following model parameters are set, $S(t_0) = 1$, $r = 0.1$, $\sigma = 0.2$. The option's maturity is $T = 1$ and strike $K = 0.95$. On a time grid stock path $S(t_i)$ is simulated. Based on these paths, we perform the hedging strategy, according to Equations (3.42) and (3.43). In Figure 3.7 three stock paths $S(t)$ are presented, for one of them the option would be in-the-money at time T (upper left), one ends out-of-money (upper right) and is at-the-money (lower left) at time T. In the three graphs $\Delta(t_i)$ (green line) behaves like the stock process $S(t)$, however when the stock $S(t)$ give a call price (pink line) deep in or out of the money, $\Delta(t_i)$ is either 0 or 1. In Table 3.1 the results for these three paths are summarized. In all three cases, the initial hedge quantities are the same, however, they change over time with the stock paths. In all three cases the final P&L(t_m) is close to zero, see Table 3.1.

Example 3.3.3 (Re-balancing frequency) A crucial assumption in the Black-Scholes model is the omission of transaction costs. This basically implies that, free-of-charge, re-balancing can be performed at any time. In practice, however, this is unrealistic. Depending on the derivative details, re-balancing

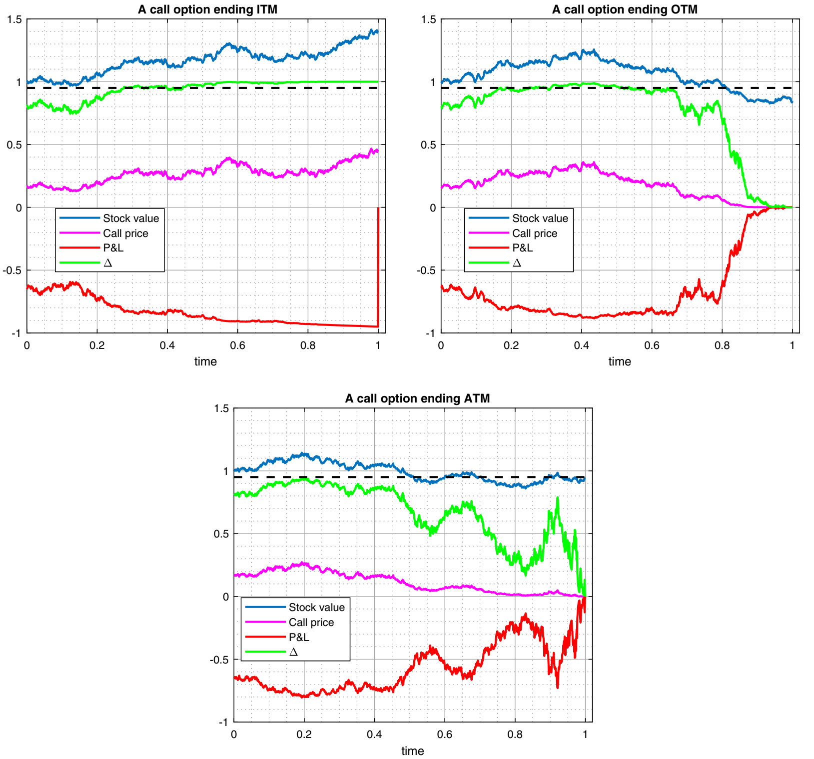


Figure 3.7: *Delta hedging a call option. Blue: the stock path, pink: the value of a call option, red: P&L(t) portfolio, and green: Δ.*

Table 3.1: *Hedge results for three stock paths.*

path no.	$S(t_0)$	P&L(t_0)	P&L(t_{m-1})	$S(t_m)$	$(S(t_m)-K)^+$	P&L(t_m)
path 1 (ITM)	1	-0.64	-0.95	1.40	0.45	$2.4 \cdot 10^{-4}$
path 2 (OTM)	1	-0.64	0.002	-0.08	0	$2.0 \cdot 10^{-4}$
path 3 (ATM)	1	-0.64	-0.010	0.96	0.01	$-2.0 \cdot 10^{-3}$

can take place daily, weekly or even on a monthly basis. When re-balancing takes place, transaction costs may need to be paid. So, for each derivative, the hedging frequency will be based on a balance between the cost and the impact of hedging.

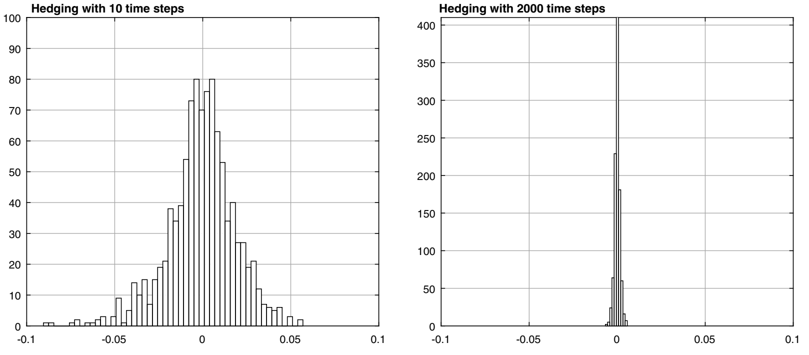


Figure 3.8: *The impact of the re-balancing frequency on the variance of the P&L(T) portfolio. Left: 10 re-balancing times, Right: 2000 re-balancing times.*

In Figure 3.8 the impact of the frequency of updating the hedge portfolio on the distribution of the P&L(T) is presented. Two simulations have been performed, one with 10 re-balancing actions during the option's life time and one with 2000 actions. It is clear that frequent re-balancing brings the variance of the portfolio P&L(T) down to almost 0.

3.4 Exercise set

Exercise 3.1 Today's asset price is $S_0 = €10$. A call on this stock with expiry in 60 days, and a strike price of $K = €10$ costs $V_c(t_0, S_0) = €1$. Suppose we bought either the stock, or we bought the call.

Fill the table below. The first row consists of possible asset prices $S(T)$, ranging from €8.5 to €11.5. Write in the second row the stock profit, in percentages, if we had bought the stock; in the third row the profit, in percentages, when having bought the option at t_0. Compare the profits and losses of having only the stock or having bought the option.

	stock price at $t = T$						
	€8.5	€9	€9.5	€10	€10.5	€11	€11.5
profit $S(T)$							
profit $V_c(T, S(T))$							

Exercise 3.2 Give at least six assumptions that form the foundation for the derivation of the Black-Scholes equation.

Exercise 3.3 Draw the payoff of a call option, as a function of strike price K, and draw in the same picture the call value at a certain time $t < T$ as a function of K.

Exercise 3.4 Assume we have a European call and put option (with the same exercise date $T = 1/4$, i.e., exercise in three months, and strike price $K = 10€$). The current stock price is 11€, and suppose a continuously compounded interest rate, $r = 6\%$. Define an arbitrage opportunity when both options are worth 2.5 €.

Exercise 3.5 Consider a portfolio Π consisting of one European option with value $V(S, t)$ and a negative amount, $-\Delta$, of the underlying stock, i.e.

$$\Pi(t) = V(S, t) - \Delta \cdot S(t).$$

With the choice $\Delta = \partial V/\partial S$, the change in this portfolio, in excess of the risk-free rate r, is given by,

$$\mathrm{d}\Pi(t) - r\Pi(t)\mathrm{d}t = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV\right)\mathrm{d}t.$$

In the Black-Scholes argument for European options, this expression should be equal to zero, since this precludes arbitrage.

Give an example of an arbitrage trade, if the equality $\mathrm{d}\Pi(t) = r\Pi(t)\mathrm{d}t$ doesn't hold, for European options, by buying and selling the portfolio $\Pi(t)$ at the time points t and $t + \Delta t$, respectively.

Exercise 3.6 Derive the European option pricing equation assuming the underlying is governed by arithmetic Brownian motion,

$$\mathrm{d}S(t) = \mu\mathrm{d}t + \sigma\mathrm{d}W^{\mathbb{P}}(t).$$

Exercise 3.7 Use a stochastic representation result (Feynman Kac theorem) to solve the following boundary value problem in the domain $[0,T] \times \mathbb{R}$.

$$\frac{\partial V}{\partial t} + \mu X \frac{\partial V}{\partial X} + \frac{1}{2}\sigma^2 X^2 \frac{\partial^2 V}{\partial X^2} = 0,$$
$$V(T,X) = \log(X^2),$$

where μ and σ are known constants.

Exercise 3.8 Consider the Black-Scholes option pricing equation,

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0.$$

a. Give boundary conditions (i.e., put values at $S = 0$ and for S "large") and the final condition (i.e. values at $t = T$) for a European put option.

b. Confirm that the expression

$$V_p(t,S) = -SF_{\mathcal{N}(0,1)}(-d_1) + Ke^{-r(T-t)}F_{\mathcal{N}(0,1)}(-d_2),$$

with

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

(plus-sign for d_1, minus-sign for d_2), with $F_{\mathcal{N}(0,1)}(\cdot)$ the standard normal CDF, satisfies the Black-Scholes equation and is indeed the solution of a put option. (You may use the two identities $F_{\mathcal{N}(0,1)}(x) + F_{\mathcal{N}(0,1)}(-x) = 1$ and

$$SF'_{\mathcal{N}(0,1)}(d_1) - e^{-r(T-t)}KF'_{\mathcal{N}(0,1)}(d_2) = 0,$$

with $F'_{\mathcal{N}(0,1)}(\cdot)$ the derivative of $F_{\mathcal{N}(0,1)}(\cdot)$.)

c. Find the solution of the European call option with the help of the put-call parity relation.

Exercise 3.9 Confirm the analytic solution

a. of the plain vanilla call option, given in Theorem 3.2.2.

b. of the digital vanilla call option in Equation (3.26).

Exercise 3.10 Consider two portfolios: π_A, consisting of a call option plus Ke^{-rT} cash; π_B, consisting of one asset S_0.

a. Based on these two portfolios, determine the following bounds for the European call option value, at time $t = 0$:

$$\max(S_0 - Ke^{-rT}, 0) \leq V_c(0, S_0) \leq S_0.$$

b. Derive the corresponding bounds for the time-zero value of the European put option $V_p(0,S_0)$, by the put-call parity relation.

c. Derive

$$\Delta = \frac{\partial V_c(t,S)}{\partial S},$$

and also the limiting behaviour of Δ, for $t \to T^-$.

c. Use again the put-call parity relation to derive the delta of a put option.

d. Give a financial argument which explains the value of $\partial V_p(t,S)/\partial S$ at expiry $t = T$, for an in-the-money option, as well as for an out-of-the-money option.

Exercise 3.11 A strangle is an option construct where an investor takes a long position in a call and in a put on the same share S with the same expiry date T, but with different strike prices (K_1 for the call, and K_2 for the put, with $K_1 > K_2$).

a. Draw the payoff of the strangle. Give an overview of the different payments possible at the expiry date. Make a distinction between three different possibilities for stock price $S(T)$.
b. Can the strange be valued with the Black-Scholes equation? Determine suitable boundary conditions, i.e., strangle values at $S = 0$ and for S "large", and final condition (value at $t = T$) for the strangle.
c. When would an investor buy a strangle with $K_2 \ll K_1$, and $K_2 < S_0 < K_1$?

Exercise 3.12 Consider the Black-Scholes equation for a cash-or-nothing call option, with the "in-the-money" value, $V^{\text{cash}}(T, S(T)) = A$, for $t = T$.

a. What are suitable boundary conditions, at $S(t) = 0$ and for $S(t)$ "large", for this option, where $0 \leq t \leq T$?
b. The analytic solution is given by $V_c^{\text{cash}}(t,S) = Ae^{-r(T-t)}F_{\mathcal{N}(0,1)}(d_2)$. Derive a put-call parity relation for the cash-or-nothing option, for all $t \leq T$, and determine the value of a cash-or-nothing put option (same parameter settings).
c. Give the value of the delta for the cash-or-nothing call, and draw in a picture the delta value just before expiry time, at $t = T^-$.

CHAPTER 4

Local Volatility Models

In this chapter:

The main problem with the Black-Scholes model is that it is not able to reproduce the *implied volatility smile*, which is present in many financial markets (see **Section 4.1**). This fact forms an important motivation to study and use *other asset models*. Alternative models are briefly discussed in Section 4.1.3. In this chapter, we give an introduction into asset prices based on *nonparametric local volatility* models. The basis for these models is a relation between (market) option prices and the implied probability density function, which we describe in **Section 4.2**. The derivation of the *nonparametric local volatility models* is found in **Section 4.3**. The pricing equation for the option values under these asset price processes appears to be a *PDE with time-dependent parameters.*

Keywords: implied volatility, smile and skew, alternative models for asset prices, implied asset density function, local volatility, arbitrage conditions, arbitrage-free interpolation.

4.1 Black-Scholes implied volatility

When the asset price is modeled by geometric Brownian motion (GBM) with constant volatility σ, the Black-Scholes model gives the unique fair value of an option contract on that underlying asset. This option price is a monotonically increasing function of the stock's volatility. High volatility means higher probability for an option to be in-the-money (ITM), see the definition in Definition 3.1.2, at the maturity date, so that the option becomes relatively more expensive.

4.1.1 The concept of implied volatility

European options in the financial markets are sometimes quoted in terms of their so-called *implied volatilities*, instead of the commonly known bid and ask option prices. In other words, the Black-Scholes implied volatility is then regarded as *a language* in which option prices can be expressed.

There are different ways to calculate the implied volatility, and we discuss a common computational technique.

For a given interest rate r, maturity T and strike price K, we have market put and call option prices on a stock S, that we will denote by $V_p^{mkt}(K,T)$, $V_c^{mkt}(K,T)$, respectively.

Focussing on calls here, the Black-Scholes implied volatility σ_{imp} is the σ-value, for which

$$V_c(t_0, S; K, T, \sigma_{\text{imp}}, r) = V_c^{mkt}(K,T), \tag{4.1}$$

where $t_0 = 0$.

The implied volatility σ_{imp} is defined as the volatility inserted as a parameter in the Black-Scholes solution that reproduces the market option price $V_c^{mkt}(K,T)$ at time $t_0 = 0$.

Whenever we speak of implied volatility in this book, we actually mean "Black-Scholes implied volatility"!

Newton-Raphson iterative method

There is no general closed-form expression for the implied volatility in terms of the option prices, so a *numerical technique* needs to be employed to determine its value. A common method for this is the Newton-Raphson root-finding iteration. The problem stated in (4.1) can be reformulated as the following root-finding problem,

$$g(\sigma_{\text{imp}}) \;:=\; V_c^{mkt}(K,T) - V_c(t_0 = 0, S_0; K, T, \sigma_{\text{imp}}, r) = 0. \tag{4.2}$$

Given an initial guess[1] for the implied volatility, i.e. $\sigma_{\text{imp}}^{(0)}$, and the derivative of $g(\sigma_{\text{imp}})$ w.r.t. σ_{imp}, we can find the next approximations, $\sigma_{\text{imp}}^{(k)}, k = 1, 2, \ldots$, by means of the *Newton-Raphson iterative process*, as follows

$$\sigma_{\text{imp}}^{(k+1)} = \sigma_{\text{imp}}^{(k)} - \frac{g(\sigma_{\text{imp}}^{(k)})}{g'(\sigma_{\text{imp}}^{(k)})}, \quad \text{for } k \geq 0. \tag{4.3}$$

Within the context of the Black-Scholes model, European call and put prices, as well as their derivatives, are known in closed-form, so that the derivative in

[1] Iterants within an iterative process are here denoted by a superscript in brackets.

Expression (4.3) can be found analytically as well, as

$$g'(\sigma) = -\frac{\partial V(t_0, S_0; K, T, \sigma, r)}{\partial \sigma} = -K\mathrm{e}^{-r(T-t_0)} f_{\mathcal{N}(0,1)}(d_2)\sqrt{T-t_0},$$

with $f_{\mathcal{N}(0,1)}(\cdot)$ the standard normal probability density function, $t_0 = 0$, and

$$d_2 = \frac{\log(\frac{S_0}{K}) + (r - \frac{1}{2}\sigma^2)(T-t_0)}{\sigma\sqrt{T-t_0}}.$$

Derivative $\frac{\partial V}{\partial \sigma}$ in $g'(\sigma)$ is, in fact, another option sensitivity, called the option's *vega*, the sensitivity with respect to volatility changes. It is another important hedge parameter. In the case where an analytic expression for the derivative is not known, additional approximations are required.

Example 4.1.1 We return to the example in Chapter 3, in Subsection 3.2.1, where the call option values and two hedge parameters, delta and gamma, were displayed for all time points and possible asset values. Here, in Figure 4.1, we will present the vega values in the (t, S)-surface, so that we see the corresponding values of the option vega at each point of the asset path followed. The following parameter set is used:

$$S_0 = 10, r = 0.05, \sigma = 0.4, T = 1, K = 10.$$

♦

Example 4.1.2 (Black-Scholes implied volatility) Consider a call option on a non-dividend paying stock $S(t)$. The current stock price is $S_0 = 100$, interest rate $r = 5\%$, the option has a maturity date in one year, struck at $K = 120$, and it is traded at $V_c^{mkt}(K, T) = 2$. The task is to find the option's implied volatility, σ_{imp}, i.e.

$$g(\sigma_{\text{imp}}) = 2 - V_c(0, 100; 120, 1y, \sigma_{\text{imp}}, 5\%) = 0.$$

The result is $\sigma_{\text{imp}} = 0.161482728841394...$. By inserting this value into the Black-Scholes equation, we obtain a call option value of 1.999999999999996.

Combined root-finding method

The Newton-Raphson algorithm converges quadratically, when the initial guess is *in the neighborhood* of a root. It is therefore important that the initial guess, $\sigma_{\text{imp}}^{(0)}$, is chosen *sufficiently close* to the root. A convergence issue is encountered with this method when the denominator in Expression (4.3) is very close to zero, or, in our context, when the option's vega is almost zero. The option vega is very small, for example, for deep ITM and OTM options, giving rise to convergence issues for the Newton-Raphson iterative method.

A basic root-finding procedure, like the bisection method may help to get a sufficiently close starting value for the Newton-Raphson iteration, or in the case

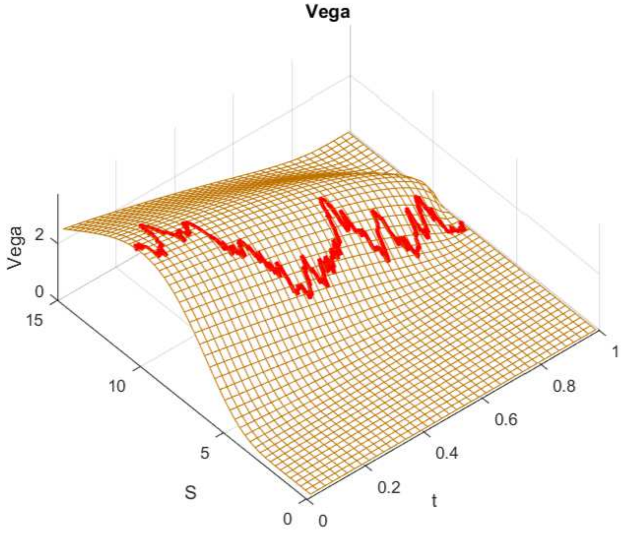


Figure 4.1: *Surface of vega values for a call with $T = 1$ in the (t, S) domain. An asset path intersects the surface so that at each $(t, S(t))$ the vega can be read.*

of serious convergence issues. Then, these two techniques form a *combined root-finding procedure.*

The method combines the efficiency of the Newton-Raphson method and the robustness of the bisection method. Similar to bisection, the combined root-finding method is based on the fact that the root should lie in an interval, here $[\sigma_l, \sigma_r]$. The method aims to find the root by means of the Newton-Raphson iteration. However, if the approximation falls outside $[\sigma_l, \sigma_r]$, the bisection method is employed to reduce the size of the interval. After some steps, the bisection method will have reduced the interval sufficiently, so that the Newton-Raphson stage will converge. The algorithm is presented below.

Combined root finding algorithm:

Given a function $g(\sigma) = 0$, find the root $\sigma = \sigma_{\text{imp}}$, with $g(\sigma_{\text{imp}}) = 0$.
Determine initial interval $[\sigma^l_{\text{imp}}, \sigma^r_{\text{imp}}]$
If $g(\sigma^l_{\text{imp}}) \cdot g(\sigma^r_{\text{imp}}) > 0$ *stop* (there is no zero in the interval)
If $g(\sigma^l_{\text{imp}}) \cdot g(\sigma^r_{\text{imp}}) < 0$ then
$k = 1$
$\sigma^{(k)}_{\text{imp}} = \frac{1}{2}(\sigma^l_{\text{imp}} + \sigma^r_{\text{imp}})$

$\delta = -g(\sigma_{\text{imp}}^{(k)})/g'(\sigma_{\text{imp}}^{(k)})$
while $\delta/\sigma_{\text{imp}}^{(k)} > \text{tol}$
  $\sigma_{\text{imp}}^{(k+1)} = \sigma_{\text{imp}}^{(k)} + \delta$
  if $\sigma_{\text{imp}}^{(k+1)} \notin [\sigma_{\text{imp}}^{l}, \sigma_{\text{imp}}^{r}]$, then
    if $g(\sigma_{\text{imp}}^{l}) \cdot g(\sigma_{\text{imp}}^{(k+1)}) > 0$, then $\sigma_{\text{imp}}^{l} = \sigma_{\text{imp}}^{(k)}$
    if $g(\sigma_{\text{imp}}^{l}) \cdot g(\sigma_{\text{imp}}^{(k+1)}) < 0$, then $\sigma_{\text{imp}}^{r} = \sigma_{\text{imp}}^{(k)}$
    $\sigma_{\text{imp}}^{(k+1)} = (\sigma_{\text{imp}}^{l} + \sigma_{\text{imp}}^{r})/2$
  $\delta = -g(\sigma_{\text{imp}}^{(k+1)})/g'(\sigma_{\text{imp}}^{(k+1)})$
  $k = k + 1$
continue

For the deep ITM and OTM options, for which the Newton-Raphson method does not converge, the combined method keeps converging, however, with a somewhat slower convergence rate. In the combined root-finding method above, derivative information is included, in the form of the option's vega in the Newton-Raphson step.

Remark 4.1.1 (Brent's method) *A* derivative-free, black-box *robust and efficient combined root-finding procedure was developed in the 1960s by van Wijngaarden, Dekker, and others [Dekker, 1969; Brent, 2013], which was subsequently improved by Brent [Brent, 1971]. The method, known as Brent's method is guaranteed to converge, so long as the function can be evaluated within an initial interval which contains a root. Brent's method combines the bisection technique with an inverse quadratic interpolation technique. The inverse quadratic interpolation technique is based on three previously computed iterates,* $\sigma^{(k)}, \sigma^{(k-1)}, \sigma^{(k-2)}$ *through which the inverse quadratic function is fitted. Iterate* $\sigma^{(k+1)}$ *is a quadratic function of* y*, and the* $\sigma^{(k+1)}$*-value at* $y = 0$ *is the next approximation for the root* σ_{imp}*.*

If the three point pairs are $[\sigma^{(k)}, g(\sigma^{(k)})], [\sigma^{(k-1)}, g(\sigma^{(k-1)})], [\sigma^{(k-2)}, g(\sigma^{(k-2)})]$*, then the interpolation formula is given by,*

$$\begin{aligned}\sigma^{(k+1)} = & \frac{g(\sigma^{(k-1)})g(\sigma^{(k-2)})\sigma^{(k)}}{(g(\sigma^{(k)}) - g(\sigma^{(k-1)}))(g(\sigma^{(k)}) - g(\sigma^{(k-2)}))} \\ & + \frac{g(\sigma^{(k-2)})g(\sigma^{(k)})\sigma^{(k-1)}}{(g(\sigma^{(k-1)}) - g(\sigma^{(k-2)}))(g(\sigma^{(k-1)}) - g(\sigma^{(k)}))} \\ & + \frac{g(\sigma^{(k-1)})g(\sigma^{(k)})\sigma^{(k-2)}}{(g(\sigma^{(k-2)}) - g(\sigma^{(k-1)}))(g(\sigma^{(k-2)}) - g(\sigma^{(k)}))}. \end{aligned} \tag{4.4}$$

When, however, that estimate lies outside the bisected interval, then it is rejected and a bisection step will take place instead.

If two consecutive approximations are identical, for example, $g(\sigma^{(k)}) = g(\sigma^{(k-1)})$*, the quadratic interpolation is replaced by the secant method,*

$$\sigma^{(k+1)} = \sigma^{(k-1)} - g(\sigma^{(k-1)}) \frac{\sigma^{(k-1)} - \sigma^{(k-2)}}{g(\sigma^{(k-1)}) - g(\sigma^{(k-2)})}. \tag{4.5}$$

▲

4.1.2 Implied volatility; implications

The volatility which is prescribed in the Black-Scholes solution is supposed to be a known constant or deterministic function of time. This is one of the basic assumptions in the Black-Scholes theory, leading to the Black-Scholes equation. This, however, is inconsistent with the observations in the financial market, because numerical inversion of the Black-Scholes equation based on market option prices, for different strikes and a fixed maturity time, exhibit a so-called *implied volatility skew or smile*. Figure 4.2 presents typical implied volatility shapes that we find in financial market data quotes. Here the implied volatility is shown against the strike dimension, varying K. Figure 4.3 shows two typical implied volatility surfaces in strike price and time dimension, where it is clear that the surface varies also in the time-wise direction. This is called the *term-structure* of the implied volatility surface.

It is without any doubt that the Black-Scholes model, and its notion of hedging option contracts by stocks and cash, forms the foundation of modern finance. Certain underlying assumptions are, however, questionable in practical market applications. To name a few, in the Black-Scholes world, delta hedging is supposed to be a continuous process, but in practice it is a discrete process (a hedged portfolio is typically updated once a week or so, depending on the type of financial derivative), and transaction costs for re-balancing of the portfolio are not taken into account. Empirical studies of financial time series have also revealed that the normality assumption for asset returns, $\mathrm{d}S(t)/S(t)$, in the Black-Scholes theory cannot capture *heavy tails* and *asymmetries*, that are found to be present in log-asset returns in the financial market quotes [Rubinstein, 1994]. Empirical densities are usually peaked compared to the normal density; a phenomenon which is known by *excess of kurtosis.*

The main problem with the Black-Scholes model is however that it is *not able to reproduce the above mentioned implied volatility skew or smile*, that is commonly observed in many financial markets. This fact forms an important motivation to study and use alternative mathematical asset models.

In this chapter, the local volatility process is discussed in detail as a first alternative model for the stock price dynamics.

4.1.3 Discussion on alternative asset price models

To overcome the issues with the Black-Scholes dynamics, a number of alternative models have been suggested in the financial literature. These asset models include local volatility models [Dupire, 1994; Derman and Kani, 1998; Coleman *et al.*,

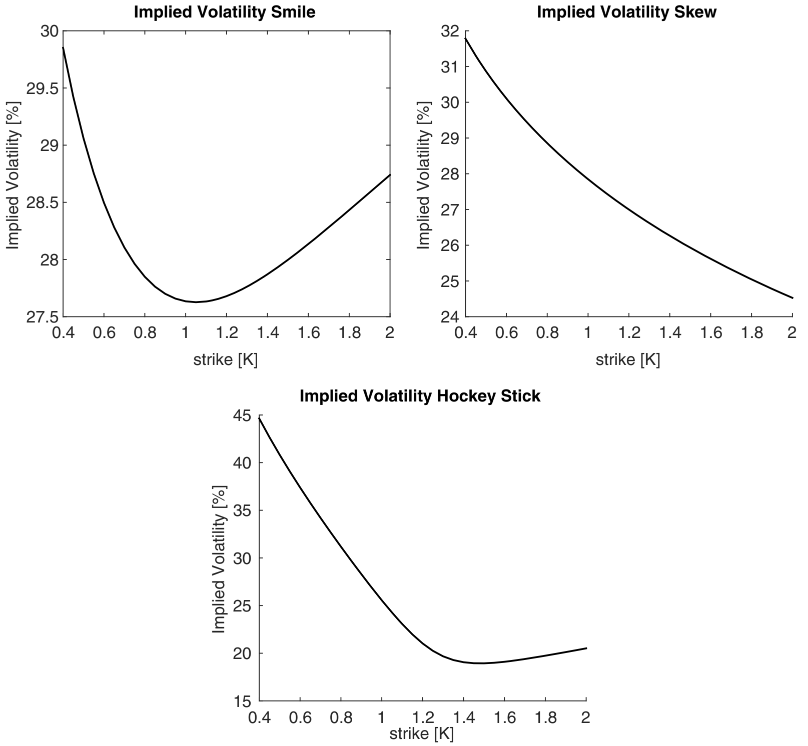


Figure 4.2: *Typical implied volatility shapes: a smile, a skew and the so-called hockey stick. The hockey stick can be seen as a combination of the implied volatility smile and the skew.*

1999; Dupire, 1994], stochastic volatility models [Hull and White, 1987; Heston, 1993], jump diffusion models [Kou, 2002; Kou and Wang, 2004; Merton, 1976] and, generally, Lévy models of finite andinfinite activity [Barndorff, 1998; Carr *et al.*, 2002; Eberlein, 2001; Matache *et al.*, 2004; Raible, 2000]. It has been shown that several of these advanced asset models are, at least to some extent, able to generate the observed market volatility skew or smile.

Although many of the above mentioned models can be fitted to the option market data, a drawback is the need for the *calibration* to determine the open parameters of the underlying stock process, so that model and market option prices fit. An exception is formed by the so-called *local volatility* (LV) models. Since the input for the LV models are the market observed implied volatility values, the LV model can be calibrated *exactly* to any given set of arbitrage-free European vanilla option prices. Local volatility models therefore do not require

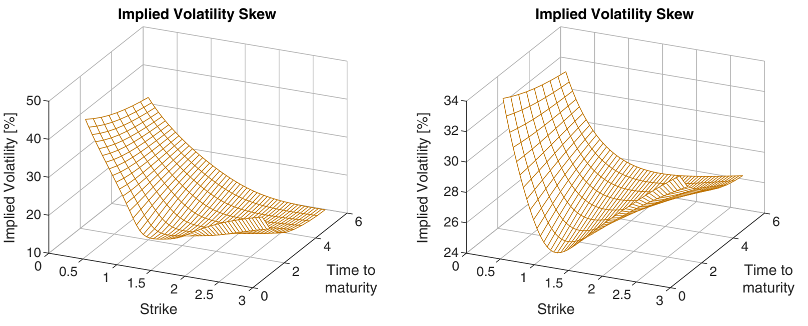


Figure 4.3: *Implied volatility surfaces. A pronounced smile for the short maturity and a pronounced skew for longer maturities T (left side figure), and a pronounced smile over all maturities (right side).*

calibration in the sense of finding the optimal model parameters, because the volatility function can be expressed in terms of the market quotes of call and/or put options. By the local volatility framework we can thus exactly reproduce market volatility smiles and skews.

For many years already, the LV model, as introduced by Dupire [1994] and Derman & Kani [1998], is considered a standard model for pricing and managing risks of financial derivative products.

Although well-accepted by practitioners for the accurate valuation of several option products, the LV model also has its limitations. For example, a *full matrix of option prices (meaning many strikes and maturity times)* is required to determine the implied volatilities. If a complete set of data is not available, as it often happens in practice, interpolation and extrapolation methods have to be applied, and they may give rise to inaccuracies in the option prices. There are other drawbacks, like an *inaccurate pricing of exotic options*, due to a local volatility model.

One may argue that an LV model does not reflect *"the dynamics of the market factors"*, because parameters are just fitted. The use of such a model outside the frame of its calibration, for different option products, different contract periods and thus volatilities, should be considered *with care*. This is an important drawback for example for the valuation of so-called forward starting options with longer maturities [Rebonato, 1999] to be covered in Chapter 10.

Example 4.1.3 Here we re-consider the two asset models from Example 2.2.1, one with a time-dependent volatility and another with a time-averaged constant volatility. In that example, we have shown that the same European option prices can be obtained by means of these two models with different asset dynamics.

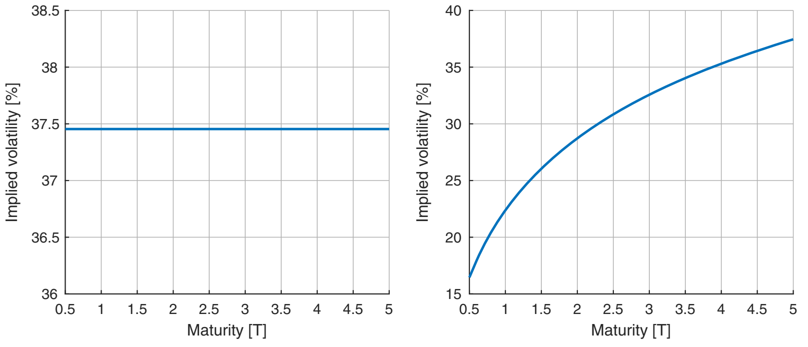


Figure 4.4: *Comparison of the volatility term structure (for ATM volatilities) for Black-Scholes model with constant volatility σ_* (left-side picture) versus a model with time-dependent volatility $\sigma(t)$ (right-side picture).*

In Figure 4.4, we now consider the implied volatility term-structure for these two models, one with a constant (*time-averaged*) volatility parameter σ_* and the other with a time-dependent volatility function $\sigma(t)$.

By means of the time-dependent volatility model, we can describe a so-called ATM volatility term structure (a time-dependent volatility structure based on at-the-money options), whereas for the model with the constant volatility model the volatility term structure is just a constant. Modeling the volatility as a time-dependent function may thus be desirable, especially in the context of hedging. Hedging costs are reduced typically by any model which explains the option pricing reality accurately. However, we typically do not know the specific time-dependent form of a volatility function beforehand. The *implied volatility* $\sigma_{\text{imp}}(T)$ may directly be obtained from a time-dependent volatility function $\sigma(t)$, via

$$\sigma_{\text{imp}}(T) = \sqrt{\frac{1}{T}\int_0^T \sigma^2(t)\mathrm{d}t},$$

by a property called *additivity of variance*. ♦

4.2 Option prices and densities

Local volatility models are based on the direct connection between option prices, implied density functions and Arrow-Debreu securities, as introduced in Section 3.2.2. The relation will become apparent in the present section.

4.2.1 Market implied volatility smile and the payoff

In this section we present a basic, but powerful, technique which was introduced by Breeden and Litzenberger [1978] for pricing European-style options with market implied volatility smile or skew patterns. The technique was actually developed 20 years before the local volatility model and it is based on the observation that the risk-neutral density for the underlying asset can be derived directly from market quotes of European call and put options. As discussed, see Equation (3.36), the calculation of the risk-neutral density from option prices requires second-order differentiation of the options with respect to the strike prices. This can be avoided with this approach as it is proposed to *differentiate the payoff function* instead of the option prices. Differentiation of the payoff is typically more stable as it does not involve any interpolation of market option prices.

Because the Breeden-Litzenberger framework is related to the marginal density of the stock S(T) at maturity time T, we cannot price any payoff function which depends on multiple maturity times. As a consequence, path-dependent option payoff functions cannot be priced in this framework.

In this section, we write in our notation the dependence on the strike price K and maturity time T explicitly in the arguments list of the option value. So, we use the notation $V(t_0, S_0; K, T)$.

Suppose we wish to price a basic European option with a well-known call or put payoff, which is also called *plain-vanilla payoff*, $H(T, S)$. As the payoff function depends only on $S(T)$, today's value is given by:

$$V(t_0, S_0; K, T) = \mathrm{e}^{-r(T-t_0)} \int_0^\infty H(T, y) f_{S(T)}(y) \mathrm{d}y, \tag{4.6}$$

with $f_{S(T)}(y) \equiv f_S(T, y; t_0, S_0)$ the risk-neutral density of the stock price process at time T; $H(T, y)$ is the payoff at time T.

The start of the derivation is to differentiate the call price $V_c(t_0, S_0; K, T)$, with payoff $\max(S - K, 0)$, with respect to strike K, as follows

$$\begin{aligned} \frac{\partial V_c(t_0, S_0; K, T)}{\partial K} &= \mathrm{e}^{-r(T-t_0)} \frac{\partial}{\partial K} \int_K^{+\infty} (y - K) f_{S(T)}(y) \mathrm{d}y \\ &= -\mathrm{e}^{-r(T-t_0)} \int_K^{+\infty} f_{S(T)}(y) \mathrm{d}y, \end{aligned} \tag{4.7}$$

so that we find, for the second derivative with respect to K:

$$\boxed{\frac{\partial^2 V_c(t_0, S_0; K, T)}{\partial K^2} = \mathrm{e}^{-r(T-t_0)} f_{S(T)}(K).} \tag{4.8}$$

As presented earlier, the *implied stock density* can thus be obtained from (market) option prices via differentiation of the call prices (see Equation (3.36)), as

$$f_{S(T)}(y) = \mathrm{e}^{r(T-t_0)} \frac{\partial^2}{\partial y^2} V_c(t_0, S_0; y, T). \tag{4.9}$$

With the help of the put-call parity (3.3), the density can also be obtained from market put option prices, i.e.,

$$\begin{aligned} f_{S(T)}(y) &= \mathrm{e}^{r(T-t_0)} \frac{\partial^2}{\partial y^2} \left(V_p(t_0, S_0; y, T) + S_0 - \mathrm{e}^{-r(T-t_0)} y \right) \\ &= \mathrm{e}^{r(T-t_0)} \frac{\partial^2 V_p(t_0, S_0; y, T)}{\partial y^2}. \end{aligned}$$

These expressions for $f_{S(T)}(y)$ can be inserted into Equation (4.6).

With two possibilities for implying the density for the stock, via calls or puts, one may wonder which options are the most appropriate. This question is related to the question which options are "liquidly available in the market". Liquidity means a lot of trading and therefore a small bid-ask spread leading to a better estimate of the stock's density. In practice, we cannot state that either calls or puts are most liquidly traded, but it is well-known that out-of-the-money (OTM) options, see the definition in Definition 3.1.2, are usually more liquid than in-the-money (ITM) options. To determine the market implied density for the stock price, one should preferably choose the OTM calls and puts.

Example 4.2.1 (Black-Scholes prices with derivative to the strike price) The Black-Scholes solution that we have used in several earlier examples is considered, with parameter values chosen as:

$$S_0 = 10, r = 0.05, \sigma = 0.4, T = 1, K = 10.$$

In this example we present the call option values as a function of t and K, so we show the function V_c in the (t, K)-plane, see Figure 4.5. Note that the call option surface looks very differently when we plot it against varying strike prices K. In this section on the local volatility model, we also often encounter the derivatives of the option values with respect to strike price K. In Figure 4.5, we therefore also present these derivatives in the (t, K) plane. Notice that indeed the second derivative of the option with respect to the strike gives us the time-dependent evolution of a density function. ♦

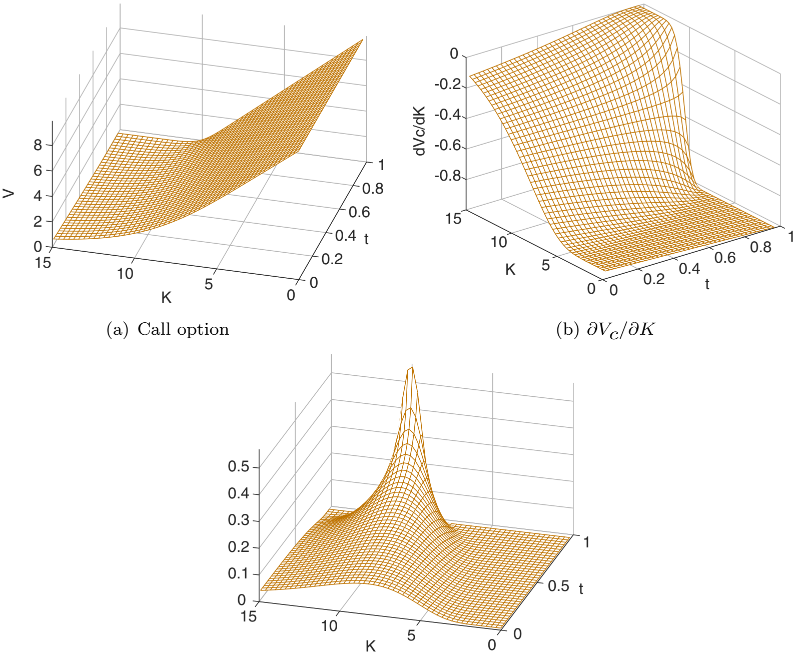
We now start by splitting the integral in (4.6) into two parts:

$$\begin{aligned} V(t_0, S_0; K, T) = \mathrm{e}^{-r(T-t_0)} \Bigg(&\int_0^{S_F} H(T, y) f_{S(T)}(y) \mathrm{d}y \\ &+ \int_{S_F}^{\infty} H(T, y) f_{S(T)}(y) \mathrm{d}y \Bigg), \end{aligned} \tag{4.10}$$

with

$$S_F \equiv S_F(t_0, T) := \mathrm{e}^{r(T-t_0)} S(t_0),$$

representing the *forward stock value*.



(a) Call option (b) $\partial V_c/\partial K$

(c) $\partial^2 V_c/\partial K^2$

Figure 4.5: *Call option values as a function of running time t and strike K, plus the first derivative with respect to strike K, i.e. $\partial V_c/\partial K$ and the corresponding second derivative $\partial^2 V_c/\partial K^2$.*

For $S > K$ we have the OTM puts and for $S < K$ we have the OTM calls, so we calculate the first integral by put prices and the second by call prices, i.e.,

$$V(t_0, S_0; K, T) = \int_0^{S_F} \underbrace{H(T, y) \frac{\partial^2 V_p(t_0, S_0; y, T)}{\partial y^2}}_{I_1(y)} \mathrm{d}y + \int_{S_F}^{\infty} \underbrace{H(T, y) \frac{\partial^2 V_c(t_0, S_0; y, T)}{\partial y^2}}_{I_2(y)} \mathrm{d}y, \tag{4.11}$$

where $\mathrm{e}^{r(T-t_0)}$ cancels out.

To obtain an accurate estimate of $V(t_0, S_0; K, T)$, we need to compute second derivatives of puts and calls at time t_0. In practice, however, only a few market quotes for these options are available, and the computed results can be sensitive to the method used to calculate the derivatives. With the help of the integration-by-parts technique, we can *exchange differentiation of the options with differentiation of the payoff function*. Assume that the payoff $H(T, y)$ is twice differentiable with respect to y. By integration by parts of $I_1(y)$ in (4.11), we find:

$$\int_0^{S_F} I_1(y)\mathrm{d}y = \left[H(T,y)\frac{\partial V_p(t_0,S_0;y,T)}{\partial y}\right]_{y=0}^{y=S_F} - \int_0^{S_F} \frac{\partial H(T,y)}{\partial y}\frac{\partial V_p(t_0,S_0;y,T)}{\partial y}\mathrm{d}y. \tag{4.12}$$

Applying again integration by parts to the last integral in (4.12), we find:

$$\int_0^{S_F} I_1(y)\mathrm{d}y = \underbrace{\left[H(T,y)\frac{\partial V_p(t_0,S_0;y,T)}{\partial y}\right]_{y=0}^{y=S_F}}_{I_{1,1}(S_F)} - \underbrace{\frac{\partial H(T,y)}{\partial y}V_p(t_0,S_0;y,T)\Big|_{y=0}^{y=S_F}}_{I_{1,2}(S_F)} + \int_0^{S_F}\frac{\partial^2 H(T,y)}{\partial y^2}V_p(t_0,S_0;y,T)\mathrm{d}y, \tag{4.13}$$

and for the second integral in (4.11), we find, similarly:

$$\int_{S_F}^{\infty} I_2(y)\mathrm{d}y = \underbrace{\left[H(T,y)\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\right]_{y=S_F}^{y=\infty}}_{I_{2,1}(S_F)} - \underbrace{\frac{\partial H(T,y)}{\partial y}V_c(t_0,S_0;y,T)\Big|_{y=S_F}^{y=\infty}}_{I_{2,2}(S_F)} + \int_{S_F}^{\infty}\frac{\partial^2 H(T,y)}{\partial y^2}V_c(t_0,S_0;y,T)\mathrm{d}y.$$

So, the option value $V(t_0, S_0; K, T)$ can be expressed as:

$$V(t_0,S_0;K,T) = I_{1,1}(S_F) - I_{1,2}(S_F) + I_{2,1}(S_F) - I_{2,2}(S_F) + \int_0^{S_F}\frac{\partial^2 H(T,y)}{\partial y^2}V_p(t_0,S_0;y,T)\mathrm{d}y + \int_{S_F}^{\infty}\frac{\partial^2 H(T,y)}{\partial y^2}V_c(t_0,S_0;y,T)\mathrm{d}y.$$

Let's have another look at the put-call parity. Differentiation of the put-call parity with respect to strike $y = K$, gives us:

$$\boxed{\frac{\partial V_c(t_0,S_0;y,T)}{\partial y} + \mathrm{e}^{-r(T-t_0)} = \frac{\partial V_p(t_0,S_0;y,T)}{\partial y}.} \tag{4.14}$$

On the other hand, by evaluating the original put-call parity (3.3) at $K = S_F \equiv S_F(t_0,T)$, we find that $V_c(t_0,S_0;K,T) + \mathrm{e}^{-r(T-t_0)}S_F = V_p(t_0,S_0;K,T) + S_0$, and therefore the ATM call and put options should satisfy the relation,

$$V_c(t_0,S_0;S_F,T) = V_p(t_0,S_0;S_F,T). \tag{4.15}$$

As a put option has value zero for strike $K = 0$, and a call option has value zero for "$K \to \infty$", we find, by (4.15),

$$\begin{aligned} I_{1,2} + I_{2,2} = & \left.\frac{\partial H(T,y)}{\partial y}\right|_{y=S_F} V_p(t_0,S_0;y,T) \\ & - \left.\frac{\partial H(T,y)}{\partial y}\right|_{y=S_F} V_c(t_0,S_0;y,T) = 0. \end{aligned} \tag{4.16}$$

With the help of (4.14), we also find:

$$\begin{aligned} I_{1,1} + I_{2,1} &= \left[H(T,y)\frac{\partial V_p(t_0,S_0;y,T)}{\partial y}\right]_{y=0}^{y=S_F} + \left[H(T,y)\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\right]_{y=S_F}^{y=\infty} \\ &= \left[H(T,y)\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\right]_{y=0}^{y=\infty} + \mathrm{e}^{-r(T-t_0)}\left(H(T,S_F) - H(T,0)\right). \end{aligned}$$

The partial derivative of a call option with respect to the strike is given by the following expression:

$$\begin{aligned} \frac{\partial V_c(t_0,S_0;K,T)}{\partial K} &= -\mathrm{e}^{-r(T-t_0)} \int_K^\infty f_{S(T)}(y)\mathrm{d}y, \\ &= -\mathrm{e}^{-r(T-t_0)}\left(1 - F_{S(T)}(K)\right), \end{aligned} \tag{4.17}$$

where $F_{S(T)}(\cdot)$ is the CDF of stock $S(T)$ at time T. This implies,

$$\begin{aligned} \left[H(T,y)\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\right]_{y=0}^{y=\infty} = & -\mathrm{e}^{-r(T-t_0)}\left[H(T,\infty)\left(1 - F_{S(T)}(\infty)\right)\right. \\ & \left. - \; H(T,0)\left(1 - F_{S(T)}(0)\right)\right]. \end{aligned}$$

Since a stock price is nonnegative, $F_{S(T)}(0) = 0$. Assuming that for $y \to \infty$, CDF $F_{S(T)}(y)$ converges faster to 1 than $H(T,y)$ converges to infinity, gives,

$$\begin{aligned} \left[H(T,y)\frac{\partial V_c(t_0,S_0;K,T)}{\partial K}\right]_{y=0}^{y=\infty} &= -\mathrm{e}^{-r(T-t_0)}\left[0 - H(T,0)\left(1-0\right)\right] \\ &= \mathrm{e}^{-r(T-t_0)}H(T,0). \end{aligned} \tag{4.18}$$

This means that the expressions for the terms $I_{1,1} + I_{2,1}$ are given by:

$$\begin{aligned} I_{1,1} + I_{2,1} &= \mathrm{e}^{-r(T-t_0)}H(T,0) + \mathrm{e}^{-r(T-t_0)}\left(H(T,S_F) - H(T,0)\right) \\ &= \mathrm{e}^{-r(T-t_0)}H(T,S_F). \end{aligned} \tag{4.19}$$

The pricing equation can thus be written as:

$$V(t_0, S_0; K, T) = \mathrm{e}^{-r(T-t_0)} H(T, S_F) + \int_0^{S_F} V_p(t_0, S_0; y, T) \frac{\partial^2 H(T, y)}{\partial y^2} \mathrm{d}y + \int_{S_F}^{\infty} V_c(t_0, S_0; y, T) \frac{\partial^2 H(T, y)}{\partial y^2} \mathrm{d}y. \tag{4.20}$$

This equation consists of two parts and allows for an intuitive interpretation, [Carr and Madan, 1998]. We encounter the *forward value of the contract*, which is known today, and there is a *volatility smile correction*, which is given in terms of two integrals. As the final expression does not involve differentiation of the call prices to the strike, this representation is more stable than the version presented in Equation (4.11).

Example 4.2.2 (European option) The Breeden-Litzenberger technique is applied here to value European calls and puts, where we expect Equation (4.20) to collapse to the standard option pricing result. For $H(T, y) = \max(y - K, 0)$, Equation (4.20) reads:

$$\begin{aligned} V(t_0, S_0; K, T) = \mathrm{e}^{-r(T-t_0)} &\max(S_F - K, 0) \\ &+ \int_0^{S_F} V_p(t_0, S_0; y, T) \frac{\partial^2 \max(y - K, 0)}{\partial y^2} \mathrm{d}y \\ &+ \int_{S_F}^{\infty} V_c(t_0, S_0; y, T) \frac{\partial^2 \max(y - K, 0)}{\partial y^2} \mathrm{d}y, \end{aligned} \tag{4.21}$$

with $S_F \equiv S_F(t_0, T) = \mathrm{e}^{r(T-t_0)} S(t_0)$. The second derivative of the maximum operator is not differentiable everywhere. Using the indicator function,

$$\max(y - K, 0) = (y - K)\mathbb{1}_{y-K>0}(y),$$

we have

$$\frac{\partial \left((y - K)\mathbb{1}_{y-K>0}(y)\right)}{\partial y} = \mathbb{1}_{y-K>0}(y) + (y - K)\delta(y - K),$$

with $\delta(y)$ the Dirac delta function, see (1.17). Differentiating twice gives,

$$\begin{aligned} \frac{\partial^2 \left((y - K)\mathbb{1}_{y-K>0}(y)\right)}{\partial y^2} &= \delta(y - K) + \delta(y - K) + (y - K)\delta'(y - K) \\ &= \delta(y - K), \end{aligned}$$

as $y\delta'(y) = -\delta(y)$. With these results, Equation (4.21) becomes, using the put-call parity from Equation (3.3),

$$\begin{aligned} V(t_0, S_0; K, T) = \mathrm{e}^{-r(T-t_0)} \max(S_F - K, 0) &+ \int_0^{\infty} V_c(t_0, S_0; y, T)\delta(y - K)\mathrm{d}y \\ &+ \int_0^{S_F} \left(\mathrm{e}^{-r(T-t_0)} y - S_0\right) \delta(y - K)\mathrm{d}y. \end{aligned}$$

The last integral in the expression can be further reduced, as follows:

$$\int_0^{S_F} \left(\mathrm{e}^{-r(T-t_0)}y - S_0\right)\delta(y-K)\mathrm{d}y = \mathrm{e}^{-r(T-t_0)}\int_0^{S_F}(y-S_F)\,\delta(y-K)\mathrm{d}y,$$

where the right-hand side integral is only nonzero for $K < S_F(t_0,T)$ and because $\int_{-\infty}^{+\infty} g(y)\delta(y-a)\mathrm{d}y = g(a)$ for any continuous function $g(y)$, we have

$$\begin{aligned}\int_0^{S_F} \left(\mathrm{e}^{-r(T-t_0)}y - S_0\right)\delta(y-K)\mathrm{d}y &= -\mathrm{e}^{-r(T-t_0)}\left(S_F - K\right)\mathbb{1}_{K<S_F}\\ &= -\mathrm{e}^{-r(T-t_0)}\max(S_F-K,0).\end{aligned}$$

So, the pricing equation yields:

$$V(t_0,S_0;K,T) = \int_0^{\infty} V_c(t_0,S_0;y,T)\delta(y-K)\mathrm{d}y \equiv V_c(t_0,S_0;K,T),$$

which concludes this derivation. ♦

4.2.2 Variance swaps

So far we discussed models in which the underlying asset, i.e. the stock price, was specified. Volatility is considered to be a quantity, which is mainly used for improving the fit to the market implied volatility smile or skew. In this subsection we make a step towards volatility modeling and concentrate on the volatility itself. Via financial products called *variance swaps*, we can actually trade the volatility, just like any other stock or commodity.

> A variance swap is a forward contract which, at maturity T, pays the difference between the realized variance and a predefined strike price (multiplied by a certain notional amount). Generally, a forward contract is a contract which is not traded at a regulated exchange but is traded directly between two parties, over-the-counter.

The two trading parties agree to buy or to sell an asset at a pre-defined time in the future at an agreed price. The realized variance, which is then "swapped" with an amount K, can be measured in different ways, since there is no formally defined market convention.

Volatility can, for example, be measured indirectly, by the continuous observation of the stock performance. A model variance swap payoff is then defined as:

$$\begin{aligned}H(T,S) &= \frac{252}{m}\sum_{i=1}^{m}\left(\log\frac{S(t_i)}{S(t_{i-1})}\right)^2 - K \qquad (4.22)\\ &=: \sigma_v^2(T) - K,\end{aligned}$$

for asset $S(t)$, with a given time-grid $t_0 < t_1 < \cdots < t_m = T$, the strike level K, and σ_v^2 is the realized variance of the stock over the life of the swap; 252 represents the number of business days in a given year. Typically, the strike K is set such that the value of the contract is initially equal to 0.

With a deterministic interest rate r, the contract value at time t_0 is given by,

$$V(t_0, S_0; K, T) = \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[\sigma_v^2(T) - K|\mathcal{F}(t_0)\right], \tag{4.23}$$

and the strike value K at which the value of the contract initially equals zero is given by,

$$\mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[\sigma_v^2(T) - K|\mathcal{F}(t_0)\right] = 0, \tag{4.24}$$

so that $K = \mathbb{E}^{\mathbb{Q}}\left[\sigma_v^2(T)|\mathcal{F}(t_0)\right]$.

The limit of the log-term in (4.22), as the time grid gets finer, i.e., $\Delta t = t_i - t_{i-1} \to 0$, is written as,

$$\log\frac{S(t_i)}{S(t_{i-1})} = \log S(t_i) - \log S(t_{i-1}) \xrightarrow{\Delta t \to 0} \mathrm{d}\log S(t),$$

With the stock $S(t)$ governed by the following stochastic process:

$$\mathrm{d}S(t) = rS(t)\mathrm{d}t + \sigma(t)S(t)\mathrm{d}W(t), \tag{4.25}$$

with constant r and a stochastic process $\sigma(t)$ for the volatility, the dynamics under the log-transformation read,

$$\mathrm{d}\log S(t) = \left(r - \frac{1}{2}\sigma^2(t)\right)\mathrm{d}t + \sigma(t)\mathrm{d}W(t). \tag{4.26}$$

The Itô table gives, to leading order, $(\mathrm{d}\log S(t))^2 = \sigma^2(t)\mathrm{d}t$, and thus

$$\int_{t_0}^{T}(\mathrm{d}\log S(t))^2 = \int_{t_0}^{T}\sigma^2(t)\mathrm{d}t. \tag{4.27}$$

The rigorous proof of the corresponding derivations can be found in [O. Barndorff-Nielsen, 2006].

With (4.25) and (4.26), we have

$$\frac{\mathrm{d}S(t)}{S(t)} - \mathrm{d}\log S(t) = \frac{1}{2}\sigma^2(t)\mathrm{d}t. \tag{4.28}$$

The term $252/m$ in (4.22) annualizes the realized variance (it returns a year percentage), and in the continuous case this is modeled by $1/(T - t_0)$.

In the continuous case, the variance swap contract can thus be written as,

$$H(T, S) = \frac{1}{T - t_0}\int_{t_0}^{T}\sigma^2(t)\mathrm{d}t - K =: \sigma_v^2(T) - K. \tag{4.29}$$

The strike K at which the swap should be traded is then given by

$$\begin{aligned}K &= \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{T-t_0}\int_{t_0}^{T}\sigma^2(t)\mathrm{d}t\Big|\mathcal{F}(t_0)\right]\\&= \frac{2}{T-t_0}\mathbb{E}^{\mathbb{Q}}\left[\int_{t_0}^{T}\frac{\mathrm{d}S(t)}{S(t)} - \mathrm{d}\log S(t)\Big|\mathcal{F}(t_0)\right] \qquad (4.30)\\&= \frac{2}{T-t_0}\mathbb{E}^{\mathbb{Q}}\left[\int_{t_0}^{T}\frac{\mathrm{d}S(t)}{S(t)}\Big|\mathcal{F}(t_0)\right] - \frac{2}{T-t_0}\mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S(t_0)}\Big|\mathcal{F}(t_0)\right],\end{aligned}$$

using (4.28). After simplifications and interchanging integration and taking the expectation, we find:

The *strike price* K for which the value of a variance swap equals zero at the inception time t_0, is given by

$$K = \frac{2}{T-t_0}\left(r(T-t_0) - \mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S(t_0)}\Big|\mathcal{F}(t_0)\right]\right), \qquad (4.31)$$

where $S(T)/S(t_0)$ represents the rate of return of the underlying stock.

Pricing equations for variance swaps

The VIX index is a known index which gives a measure for the implied volatility of the S&P 500 index. Its value is based on options on the S&P 500 index for which the time to maturity ranges from 23 to 37 days. An average of the implied volatility is calculated for 30 days options on the S&P500. The VIX index is also called the *fear index*, as it represents the market's expectation of the stock market volatility. A European counterpart to the VIX index is the VSTOXX Volatility index, where the underlying index is the Euro Stoxx 50 index.

We extract call and put market option prices on the S&P500. With the derivations from Section 4.2.1, we can derive pricing equations for the variance swaps. To determine strike K, we calculate the expectation in (4.31), with the help of these market option quotes.

To price the following contract

$$V(t_0,S_0;K,T) = \mathrm{e}^{-r(T-t_0)}\mathbb{E}^{\mathbb{Q}}\left[H(T,S)\Big|\mathcal{F}(t_0)\right], \qquad (4.32)$$

with some payoff function $H(T,S)$, taking into account the volatility smile, the price can be determined via Equation (4.20), i.e.

$$\begin{aligned}V(t_0,S_0;K,T) = \mathrm{e}^{-r(T-t_0)}H(T,S_F)&\\+\int_0^{S_F}V_p(t_0,S_0;y,T)\frac{\partial^2 H(T,y)}{\partial y^2}\mathrm{d}y&\\+\int_{S_F}^{\infty}V_c(t_0,S_0;y,T)\frac{\partial^2 H(T,y)}{\partial y^2}\mathrm{d}y,& \qquad (4.33)\end{aligned}$$

with the forward stock, $S_F \equiv S_F(t_0,T) := S(t_0)\mathrm{e}^{r(T-t_0)}$, and where $V_c(t_0,S_0;y,T)$ and $V_p(t_0,S_0;y,T)$ are call and put option prices, respectively.

Strike price K at which a variance swap is traded, is given by:

$$K = \frac{2}{T-t_0}\left\{r(T-t_0) - \mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S(t_0)}\Big|\mathcal{F}(t_0)\right]\right\}. \tag{4.34}$$

The expression in (4.34) can be modified by means of the forward stock,

$$\begin{aligned}K &= -\frac{2}{T-t_0}\mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S(t_0)} - \log\mathrm{e}^{r(T-t_0)}\Big|\mathcal{F}(t_0)\right]\\ &= -\frac{2}{T-t_0}\mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S_F}\Big|\mathcal{F}(t_0)\right].\end{aligned}$$

With $H(T,y) = \log\frac{y}{S_F}$, we find

$$\frac{\partial}{\partial y}H(T,y) = \frac{1}{y}, \qquad \frac{\partial^2}{\partial y^2}H(T,y) = -\frac{1}{y^2},$$

and the expectation in (4.34) can be calculated by Equations (4.32) and (4.33), as follows:

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S_F}\Big|\mathcal{F}(t_0)\right] &= \mathrm{e}^{r(T-t_0)}V(t_0,S_0;K,T)\\ &= \log\frac{S_F}{S_F} - \mathrm{e}^{r(T-t_0)}\int_0^{S_F}\frac{1}{y^2}V_p(t_0,S_0;y,T)\mathrm{d}y\\ &\quad - \mathrm{e}^{r(T-t_0)}\int_{S_F}^{\infty}\frac{1}{y^2}V_c(t_0,S_0;y,T)\mathrm{d}y\\ &= -\mathrm{e}^{r(T-t_0)}\left[\int_0^{S_F}\frac{1}{y^2}V_p(t_0,S_0;y,T)\mathrm{d}y\right.\\ &\quad \left. + \int_{S_F}^{\infty}\frac{1}{y^2}V_c(t_0,S_0;y,T)\mathrm{d}y\right],\end{aligned}$$

so that,

$$\boxed{\begin{aligned}K &= -\frac{2}{T-t_0}\mathbb{E}^{\mathbb{Q}}\left[\log\frac{S(T)}{S_F}\Big|\mathcal{F}(t_0)\right] \qquad (4.35)\\ &= \frac{2}{T-t_0}\mathrm{e}^{r(T-t_0)}\left[\int_0^{S_F}\frac{1}{y^2}V_p(t_0,S_0;y,T)\mathrm{d}y + \int_{S_F}^{\infty}\frac{1}{y^2}V_c(t_0,S_0;y,T)\mathrm{d}y\right].\end{aligned}}$$

This expression can be efficiently calculated by, for example, a Gauss quadrature integration technique. The obtained results enable us to calculate the value of a variance swap for given market quotes of European call and put option prices. Once a market quote for a variance swap differs significantly from the value implied by the market call and put options, given in (4.35), we may find a volatility arbitrage possibility.

Remark 4.2.1 ($\mathbb{E}^{\mathbb{Q}}[\log S(T)|\mathcal{F}(t_0)]$ in the Black-Scholes model)
One may argue that $\mathbb{E}^{\mathbb{Q}}\left[\log S(T)|\mathcal{F}(t_0)\right]$ *can be easily determined under the Black-Scholes model, as* $\log S(T)$ *is given by,*

$$\log S(T) = \left(r - \frac{1}{2}\sigma^2\right)(T - t_0) + \sigma\left(W(T) - W(t_0)\right),$$

and the expectation is given by,

$$\mathbb{E}^{\mathbb{Q}}\left[\log S(T)|\mathcal{F}(t_0)\right] = \left(r - \frac{1}{2}\sigma^2\right)(T - t_0),$$

which is simply expressed in terms of the interest rate r *and the volatility* σ*. However, it is unclear which volatility* σ *to use for the evaluation. A crucial difference with European option valuation is that these products (European options) are quoted in the financial market, so that the implied volatility can be determined. For* $\mathbb{E}[\log S(T)]$ *the situation is complicated as there are no liquid products that can be used to determine the volatility* σ*.*

A solution to this problem is to use "all available market quotes" to derive the expectation, which is the idea behind the Breeden-Litzenberger method. ▲

Relation of variance swaps to VIX index

With the help of the Breeden-Litzenberger methodology, the equation in (4.35) provides a direct link between variance swaps and implied distributions. A key element of the equation is that the methodology is *essentially model-free*. The integrals are only based on market quotes for call and put options.

We will connect (4.35) to the VIX volatility index, and start with some strike,[2] $K_f < S_F(t_0, T)$, for which we decompose the integrals in (4.35), as follows:

$$K = \frac{2}{T - t_0}\mathrm{e}^{r(T-t_0)}\left[\int_0^{K_f}\frac{1}{y^2}V_p(t_0, S_0; y, T)\mathrm{d}y + \int_{K_f}^{\infty}\frac{1}{y^2}V_c(t_0, S_0; y, T)\mathrm{d}y\right]$$
$$+ \frac{2}{T - t_0}\mathrm{e}^{r(T-t_0)}\int_{K_f}^{S_F}\frac{1}{y^2}\left(V_p(t_0, S_0; y, T) - V_c(t_0, S_0; y, T)\right)\mathrm{d}y. \qquad (4.36)$$

By the put-call parity, for $y = K$, we have:

$$V_p(t_0, S_0; y, T) - V_c(t_0, S_0; y, T) = \mathrm{e}^{-r(T-t_0)}y - S_0,$$

[2]Strike K_f should be the highest strike which is below the forward S_F.

so that for the last expression in (4.36), we write,

$$\begin{aligned}
&\frac{2}{T-t_0}\mathrm{e}^{r(T-t_0)} \times \int_{K_f}^{S_F} \frac{1}{y^2}(\mathrm{e}^{-r(T-t_0)}y - S_0)\mathrm{d}y \\
&\quad = \frac{2}{T-t_0}\mathrm{e}^{r(T-t_0)}\left(\mathrm{e}^{-r(T-t_0)}\left(\log S_F - \log K_f\right) + S(t_0)\left(\frac{1}{S_F} - \frac{1}{K_f}\right)\right) \\
&\quad = \frac{2}{T-t_0}\left(\log\frac{S_F}{K_f} + \left(1 - \frac{S_F}{K_f}\right)\right). \qquad (4.37)
\end{aligned}$$

Application of a Taylor series expansion to the logarithm, ignoring terms higher than second-order, results in,

$$\log\frac{S_F}{K_f} \approx \left(\frac{S_F}{K_f} - 1\right) - \frac{1}{2}\left(\frac{S_F}{K_f} - 1\right)^2. \qquad (4.38)$$

With this, the expression in (4.37) becomes

$$\begin{aligned}
&\frac{2}{T-t_0}\mathrm{e}^{r(T-t_0)} \int_{K_f}^{S_F} \frac{1}{y^2}\left(V_p(t_0, S_0; y, T) - V_c(t_0, S_0; y, T)\right)\mathrm{d}y \\
&\quad \approx -\frac{1}{T-t_0}\left(\frac{S_F}{K_f} - 1\right)^2.
\end{aligned}$$

and thus (4.36) reads:

$$\begin{aligned}
K \approx{}& \frac{2}{T-t_0}\mathrm{e}^{r(T-t_0)}\left[\int_0^{K_f} \frac{1}{y^2}V_p(t_0, S_0; y, T)\mathrm{d}y + \int_{K_f}^{\infty} \frac{1}{y^2}V_c(t_0, S_0; y, T)\mathrm{d}y\right] \\
&- \frac{1}{T-t_0}\left(\frac{S_F}{K_f} - 1\right)^2.
\end{aligned}$$

Discretizing the integrals above, in the usual way discretizing the K dimension with N_K increments ΔK, gives,

$$\begin{aligned}
K \approx{}& \frac{2}{T-t_0}\mathrm{e}^{r(T-t_0)}\left[\sum_{i=1}^{f-1} \frac{1}{K_i^2}V_p(t_0, S_0; K_i, T)\Delta K + \sum_{i=f}^{N_K} \frac{1}{K_i^2}V_c(t_0, S_0; K_i, T)\Delta K\right] \\
&- \frac{1}{T-t_0}\left(\frac{S_F}{K_f} - 1\right)^2.
\end{aligned}$$

This expression can be recognized as the square of the VIX index (denoted by VIX^2), which was defined by the Chicago Board Options Exchange (CBOE) in their white paper [CBOE White Paper], as follows,

$$\text{VIX}^2 = \frac{2}{T-t_0}\sum_{i=1}^{N_K}\frac{\Delta K}{K_i^2}\mathrm{e}^{r(T-t_0)}Q(K_i) - \frac{1}{T-t_0}\left[\frac{S_F}{K_f}-1\right]^2. \tag{4.39}$$

Here, $Q(K_i)$ represents the price of out-of-the money call and put options with strike prices K_i and K_f being the highest strike below the forward price $S_F = S_F(t_0,T) := S_0\mathrm{e}^{r(T-t_0)}$.

4.3 Non-parametric local volatility models

Although the Breeden-Litzenberger technique is useful, its applicability is essentially restricted to *plain vanilla*, European put and call options. Local volatility models, which have a wider range of applicability, are discussed next.

We consider the pricing of a call option under the *local volatility* model, based on a one-dimensional stochastic stock process. The term "local volatility" is used to indicate that the volatility of the process is a function of stock $S(t)$. The stock is not driven by an additional stochastic process, which is the case with the *stochastic volatility models*, where volatility is described by an additional stochastic process, as will be discussed in Chapter 8.

The classical local volatility model is given by

$$\mathrm{d}S(t) = rS(t)\mathrm{d}t + \sigma_{LV}(t,S)S(t)\mathrm{d}W(t), \quad S(t_0) = S_0, \tag{4.40}$$

with a constant (or deterministic) interest rate r.
For a given option value $V(t,S)$, the no-arbitrage assumption gives rise to the following PDE, using Itô's lemma, as described in Chapter 2,

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}\sigma_{LV}^2(t,S)S^2\dfrac{\partial^2 V}{\partial S^2} + rS\dfrac{\partial V}{\partial S} - rV = 0, \\ V(T,S) = \max(S(T)-K,0). \end{cases} \tag{4.41}$$

PDE (4.41) is a *backward Kolmogorov partial differential equation.*

We may describe the evolution of a probability density function (PDF) by means of a Fokker-Planck PDE, which describes the *forward evolution of a PDF* in time. In that case, with an initial condition at t_0, the PDF is given by a Dirac delta function at time t_0.

Theorem 4.3.1 (Fokker-Planck PDE and SDEs) *The transition density $f_{S(t)}(y) \equiv f_S(t, y; t_0, S_0)$ associated to the general SDE for $S(t)$, $t_0 \leq t \leq T$,*

$$\mathrm{d}S(t) = \bar{\mu}(t, S)dt + \bar{\sigma}(t, S)\mathrm{d}W(t), \quad S(t_0) = S_0,$$

satisfies the Fokker-Planck (also called "forward Kolmogorov") PDE:

$$\begin{cases} \dfrac{\partial}{\partial t} f_{S(t)}(y) + \dfrac{\partial}{\partial y}[\bar{\mu}(t, y) f_{S(t)}(y)] - \dfrac{1}{2}\dfrac{\partial^2}{\partial y^2}[\bar{\sigma}^2(t, y) f_{S(t)}(y)] = 0, \\ f_{S(t_0)}(y) = \delta(y = S_0). \end{cases} \tag{4.42}$$

Local volatility modeling is based on the Fokker-Planck forward equation. In a local volatility framework the stock density is again directly connected to the market quotes of call and put options. After some calculations, the state-dependent volatility function, $\bar{\sigma}(t, S) = S \cdot \sigma_{LV}(t, S)$, can be described by means of financial market quotes.

Unfortunately, as only a limited number of options is quoted in the market for a particular stock, it is difficult to accurately recover the stock density from these few market quotes (one would require an infinite number of options in the "strike dimension" $K \in [0, \infty)$). We will discuss important conditions related to arbitrage-free option values and the related densities, to improve the accuracy of the approximation in Section 4.3.1.

The derivative of an option value with respect to maturity time T, is given by:

$$\begin{aligned} \frac{\partial V_c(t_0, S_0; K, T)}{\partial T} &= \frac{\partial}{\partial T}\left(\mathrm{e}^{-r(T-t_0)} \int_K^{+\infty} (y - K) f_{S(T)}(y)\mathrm{d}y\right) \\ &= -rV_c(t_0, S_0; K, T) + \mathrm{e}^{-r(T-t_0)} \int_K^{+\infty} (y - K) \frac{\partial f_{S(T)}(y)}{\partial T}\mathrm{d}y. \end{aligned} \tag{4.43}$$

For the partial derivative in the last integral in (4.43), we employ the Fokker-Planck equation (4.42), with $\bar{\mu}(t, S(t)) = rS$ and $\bar{\sigma}(t, S) = \sigma_{LV}(t, S) \cdot S$, by which the integral can be written as

$$\begin{aligned} \int_K^{+\infty} (y - K) \frac{\partial f_{S(T)}(y)}{\partial T}\mathrm{d}y &= -r \int_K^{+\infty} (y - K) \frac{\partial \left(y f_{S(T)}(y)\right)}{\partial y}\mathrm{d}y \\ &\quad + \frac{1}{2} \int_K^{+\infty} (y - K) \frac{\partial^2 \left(\sigma_{LV}^2(T, y) y^2 f_{S(T)}(y)\right)}{\partial y^2}\mathrm{d}y. \end{aligned} \tag{4.44}$$

The first integral at the right-hand side is equal to[3]:

$$\begin{aligned} \int_K^{+\infty} (y - K) \frac{\partial \left(y f_{S(T)}(y)\right)}{\partial y}\mathrm{d}y &= (y - K) y f_{S(T)}(y)\Big|_{y=K}^{+\infty} - \int_K^{+\infty} y f_{S(T)}(y)\mathrm{d}y \\ &= -\int_K^{+\infty} y f_{S(T)}(y)\mathrm{d}y. \end{aligned}$$

[3]Assuming that for $y \to +\infty$ the density $f_{S(T)}(y)$ decays *faster* to zero than y^2 grows to infinity.

By Equations (4.8) and (4.17), this integral can be expressed in terms of call option values, i.e.

$$
\begin{aligned}
-\int_K^{+\infty} y f_{S(T)}(y)\mathrm{d}y &= -\mathrm{e}^{r(T-t_0)}\int_K^{+\infty} y\frac{\partial^2 V_c(t_0,S_0;y,T)}{\partial y^2}\mathrm{d}y \\
&= -\mathrm{e}^{r(T-t_0)}\left[y\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\bigg|_K^{+\infty}\right. \\
&\qquad \left. - \int_K^{+\infty}\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\mathrm{d}y\right] \\
&= -\mathrm{e}^{r(T-t_0)}\left[-y\mathrm{e}^{-r(T-t_0)}\left(1-F_{S(T)}(y)\right)\bigg|_K^{+\infty}\right. \\
&\qquad \left. - \int_K^{+\infty}\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\mathrm{d}y\right],
\end{aligned}
\tag{4.45}
$$

since $F_{S(T)}(+\infty) = 1$ and under the assumption that $F_{S(T)}(y)$ converges to 1 faster than y to $+\infty$, we have

$$
y\mathrm{e}^{-r(T-t_0)}\left(1-F_{S(T)}(y)\right)\bigg|_K^{+\infty} = -K\mathrm{e}^{-r(T-t_0)}\left(1-F_{S(T)}(K)\right). \tag{4.46}
$$

Call option prices converge to 0 for $K \to \infty$, so that

$$
\int_K^{+\infty}\frac{\partial V_c(t_0,S_0;y,T)}{\partial y}\mathrm{d}y = -V_c(t_0,S_0;K,T). \tag{4.47}
$$

By Equation (4.17), this gives,

$$
\begin{aligned}
-\int_K^{+\infty} y f_{S(T)}(y)\mathrm{d}y &= -\mathrm{e}^{r(T-t_0)}K\mathrm{e}^{-r(T-t_0)}\left(-\mathrm{e}^{r(T-t_0)}\frac{\partial V_c(t_0,S_0;K,T)}{\partial K}\right) \\
&\quad + \mathrm{e}^{r(T-t_0)}V_c(t_0,S_0;K,T) \\
&= \mathrm{e}^{r(T-t_0)}\left[K\frac{\partial V_c(t_0,S_0;K,T)}{\partial K} - V_c(t_0,S_0;K,T)\right].
\end{aligned}
$$

In a similar manner, the second integral in (4.44) can be determined, as

$$
\begin{aligned}
&\int_K^{+\infty}(y-K)\frac{\partial^2\left(\sigma^2_{LV}(T,y)y^2 f_{S(T)}(y)\right)}{\partial y^2}\mathrm{d}y \\
&= \sigma^2_{LV}(T,K)K^2 f_{S(T)}(K) \\
&= \mathrm{e}^{r(T-t_0)}\sigma^2_{LV}(T,K)K^2\frac{\partial^2 V_c(t_0,S_0;K,T)}{\partial K^2}.
\end{aligned}
$$

Collecting all terms obtained for (4.43), we find:

$$
\frac{\partial}{\partial T}V_c = -rV_c - rK\frac{\partial V_c}{\partial K} + rV_c + \frac{1}{2}\sigma^2_{LV}(T,K)K^2\frac{\partial^2 V_c}{\partial K^2}. \tag{4.48}
$$

The expression for the local volatility function $\sigma_{LV}(T, K)$ can be found from,

$$\boxed{\sigma^2_{LV}(T,K) = \frac{\dfrac{\partial V_c(t_0,S_0;K,T)}{\partial T} + rK\dfrac{\partial V_c(t_0,S_0;K,T)}{\partial K}}{\dfrac{1}{2}K^2\dfrac{\partial^2 V_c(t_0,S_0;K,T)}{\partial K^2}}.} \tag{4.49}$$

Here $V_c(t_0, S_0; K, T)$ is the representation of the market call option prices at time t_0, initial stock value S_0, strike price K and maturity T.

In the local volatility model, the volatility $\sigma_{LV}(T, K)$ can thus be described by market quotes of option values, and the local volatility model may therefore *perfectly fit* to the market option quotes while the model stays parameter-free. In other words, there is no need for a calibration procedure.

4.3.1 Implied volatility representation of local volatility

From Equation (4.49) it is clear that the local volatility $\sigma_{LV}(T, K)$ can be expressed in terms of European option prices. An option price surface is thus needed to compute this so-called *Dupire's local volatility term.* Based the local volatility model, we may subsequently price options, for example, by means of the finite difference method for the resulting option valuation PDE.

In practice, however, not all derivatives in (4.49) can be obtained directly from the market option quotes and the derivatives may also need to be calculated from the market data by finite difference approximations. For each local volatility term, we need four option values for the finite difference approximations of the partial derivatives.

In the denominator of (4.49) we encounter an approximation of the density of the stock price, see Equation (4.9), for which finite difference approximations may lead to significant approximation errors for high strike prices. The errors may be particularly large when the stock density is close to 0. As a consequence, the local volatility function $\sigma_{LV}(T, K)$ will then also be unrealistically large.

Remark 4.3.1 *A number of improvements for the local volatility's accuracy have been presented in the literature. One approach is to parameterize the option surface by a certain two-dimensional (polynomial) function, $h(T, K)$, and fit it to market data $V_c^{mkt}(K, T)$. In this case, the partial derivatives with respect to time and strike price of $h(T, K)$ can be found analytically. The main problem of this approach is that it is difficult to determine a parametric function $h(T, K)$ which fits well to all market quotes while preserving the no-arbitrage assumptions.*

An issue when fitting a polynomial to option prices is that for OTM options small differences in prices may cause significant differences in the implied volatilities, due to the Black-Scholes inversion procedure. ▲

It is useful to *reformulate the LV model* in terms of the corresponding implied volatilities. However, the partial derivatives of the call prices with respect to the strike values are not equal to the partial derivatives of the implied volatilities.

We therefore derive the local volatility $\sigma_{LV}(T,K)$ in terms of the implied (Black-Scholes) volatilities, $\sigma_{\text{imp}}(T,K)$, which indicates[4] the implied volatility at time T and strike K. We know, from Equation (3.23), that arbitrage-free European call option prices are given by:

$$V_c(t_0,S_0;K,T) = S_0 F_{\mathcal{N}(0,1)}(d_1) - K\mathrm{e}^{-r(T-t_0)} F_{\mathcal{N}(0,1)}(d_2),$$

with

$$d_1 = \frac{\log\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2_{\text{imp}}(T,K)\right)(T-t_0)}{\sigma_{\text{imp}}(T,K)\sqrt{T-t_0}}, \quad d_2 = d_1 - \sigma_{\text{imp}}(T,K)\sqrt{T-t_0},$$

We will use,

$$y := \log\left(\frac{K}{S_F}\right) = \log\left(\frac{K}{S_0}\right) - r(T-t_0), \quad w := \sigma^2_{\text{imp}}(T,K)(T-t_0), \tag{4.50}$$

where S_F is again the forward price, defined as $S_F = S_F(t_0,T) := S_0\mathrm{e}^{r(T-t_0)}$.

For the variables y and w, we define the call price, $c(y,w)$, by

$$\boxed{V_c(t_0,S_0;K,T) = S_0\Big[F_{\mathcal{N}(0,1)}(d_1) - \mathrm{e}^y F_{\mathcal{N}(0,1)}(d_2)\Big] =: c(y,w),} \tag{4.51}$$

where $d_1 = \frac{1}{2}\sqrt{w} - \frac{y}{\sqrt{w}}$, and $d_2 = d_1 - \sqrt{w}$.

Using the short-hand notation $V_c := V_c(t_0,S_0;K,T)$ and $c := c(y,w)$, and $\tau := T - t_0$, we obtain,

$$\frac{\partial V_c}{\partial K} = \frac{\partial c}{\partial y}\frac{1}{K} + \frac{\partial c}{\partial w}\frac{\partial w}{\partial K}, \tag{4.52}$$

thus

$$\frac{\partial^2 V_c}{\partial K^2} = \frac{1}{K^2}\left(\frac{\partial^2 c}{\partial y^2} - \frac{\partial c}{\partial y}\right) + \frac{2}{K}\frac{\partial w}{\partial K}\frac{\partial^2 c}{\partial w \partial y} + \frac{\partial^2 w}{\partial K^2}\frac{\partial c}{\partial w} + \left(\frac{\partial w}{\partial K}\right)^2\frac{\partial^2 c}{\partial w^2}, \tag{4.53}$$

and,

$$\frac{\partial V_c}{\partial T} = -r\frac{\partial c}{\partial y} + \frac{\partial c}{\partial w}\frac{\partial w}{\partial T}. \tag{4.54}$$

After substituting these derivatives in the local volatility, Equation (4.49) is given by:

$$\sigma^2_{LV}(T,K) = \frac{\frac{\partial c}{\partial w}\frac{\partial w}{\partial T} + rK\frac{\partial c}{\partial w}\frac{\partial w}{\partial K}}{\frac{1}{2}\left(\frac{\partial^2 c}{\partial y^2} - \frac{\partial c}{\partial y}\right) + K\frac{\partial w}{\partial K}\frac{\partial^2 c}{\partial w \partial y} + \frac{1}{2}K^2\left[\frac{\partial^2 w}{\partial K^2}\frac{\partial c}{\partial w} + \left(\frac{\partial w}{\partial K}\right)^2\frac{\partial^2 c}{\partial w^2}\right]}. \tag{4.55}$$

[4]We include here the additional arguments T and K to indicate the dependence of the implied volatility on the maturity and the strike price.

The expression above can be simplified by the application of the following identities:

$$\frac{\partial^2 c}{\partial w^2} = \frac{\partial c}{\partial w}\left(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2}\right), \quad \frac{\partial^2 c}{\partial w \partial y} = \frac{\partial c}{\partial w}\left(\frac{1}{2} - \frac{y}{w}\right), \quad \frac{\partial^2 c}{\partial y^2} = \frac{\partial c}{\partial y} + 2\frac{\partial c}{\partial w}.$$

The proofs for these equalities are based on standard derivations.

Local volatility expression (4.55) now becomes:

$$\sigma^2_{LV}(T,K) = \frac{\frac{\partial w}{\partial T} + rK\frac{\partial w}{\partial K}}{1 + K\frac{\partial w}{\partial K}\left(\frac{1}{2} - \frac{y}{w}\right) + \frac{1}{2}K^2\frac{\partial^2 w}{\partial K^2} + \frac{1}{2}K^2\left(\frac{\partial w}{\partial K}\right)^2\left(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2}\right)}. \tag{4.56}$$

We may use Equation (4.50) to determine the remaining derivatives, which results in a relation between function w and implied volatility $\sigma_{\text{imp}} := \sigma_{\text{imp}}(T,k)$:

$$\begin{aligned}
\frac{\partial w}{\partial T} &= \sigma^2_{\text{imp}} + 2(T - t_0)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial T}, \\
\frac{\partial w}{\partial K} &= 2(T - t_0)\sigma_{\text{imp}}\frac{\partial \sigma_{\text{imp}}}{\partial K}, \\
\frac{\partial^2 w}{\partial K^2} &= 2(T - t_0)\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + 2(T - t_0)\sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}.
\end{aligned}$$

Local volatility function $\sigma_{LV}(T,K)$ can thus be expressed in terms of the implied volatilities σ_{imp}. In essence, the leading term in the denominator of Equation (4.56) is a first derivative, whereas in Equation (4.49) it is a second derivative term. Generally first derivatives can be computed numerically in a more stable way than second derivatives.

An example is presented in Section 4.3.4.

4.3.2 Arbitrage-free conditions for option prices

As already mentioned, in financial markets there are usually not enough market option quotes to approximate all sorts of financial derivatives, with different strike prices and maturity times, accurately, and interpolation and extrapolation of market option quotes often has to take place. In particular, an *arbitrage-free interpolation* between the available market data needs to be introduced. Arbitrage can occur in the time as well as in the strike direction, see an (t, S_t)-surface of option values in Figure 4.6.

An interpolation between different maturities and strike prices should satisfy the following conditions:

1. The so-called *calendar spread* condition:

$$V_c(t_0, S_0; K, T + \Delta T) - V_c(t_0, S_0; K, T) > 0.$$

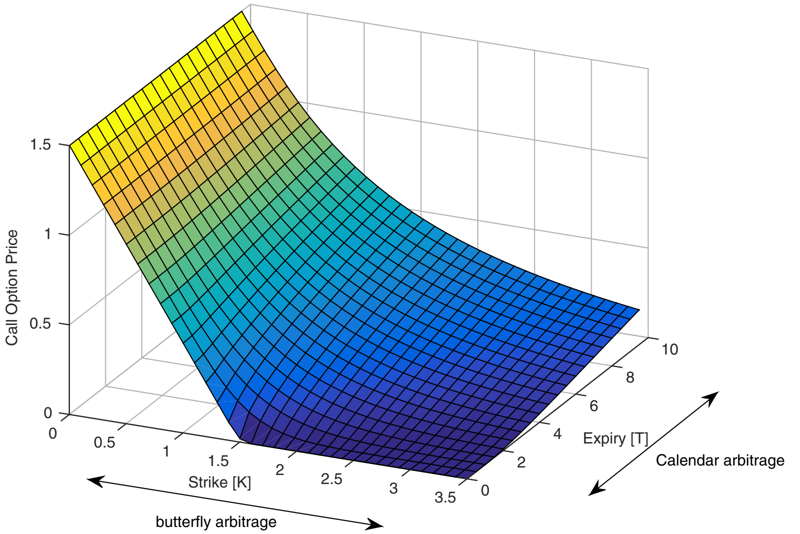


Figure 4.6: *Call prices and the two directions in which arbitrage may take place.*

If we divide both sides of the equation by ΔT and we let $\Delta T \to 0$, we find:

$$\lim_{\Delta T \to 0} \frac{1}{\Delta T} \left[V_c(t_0, S_0; K, T + \Delta T) - V_c(t_0, S_0; K, T) \right]$$
$$= \frac{\partial}{\partial T} V_c(t_0, S_0; K, T).$$

This condition can be interpreted as follows: for any two European-style options that have the same contract details, apart from the expiry date T, the contract which expires first has to be cheaper. This property is related to the fact that an option with a larger time to expiry, $T - t_0$, has a higher chance to get in-the-money.

2. *Monotonicity in the strike direction*:

$$V_c(t_0, S_0; K + \Delta K, T) - V_c(t_0, S_0; K, T) < 0, \quad \text{for calls,}$$
$$V_p(t_0, S_0; K + \Delta K, T) - V_p(t_0, S_0; K, T) > 0, \quad \text{for puts.}$$

As before in the limit we have:

$$\lim_{\Delta K \to 0} \frac{1}{\Delta K} \left(V_c(t_0, S_0; K + \Delta K, T) - V_c(t_0, S_0; K, T) \right)$$
$$= \frac{\partial}{\partial K} V_c(t_0, S_0; K, T).$$

In other words, *the most expensive* call option is the one with strike price $K = 0$. Since the call payoff value, $\max(S(T) - K, 0)$, is monotonically decreasing in strike K, any European call option with $K \neq 0$ has to be decreasing in price with decreasing K. The opposite can be concluded for put options, by using the call-put parity relation.

3. The so-called *butterfly condition* states,

$$V_c(t_0, S_0; K + \Delta K, T) - 2V_c(t_0, S_0; K, T) + V_c(t_0, S_0; K - \Delta K, T) \geq 0. \tag{4.57}$$

The nonnegativity of the expression above is understood from the following argument. If we buy one call option with strike $K + \Delta K$, sell two options with strike K and buy one call option with strike $K - \Delta K$, the resulting payoff cannot be negative (excluding transaction costs). Since any discounted nonnegative value stays nonnegative, Equation (4.57) needs to hold. As explained in Equation (3.36), the second derivative of the call option price can be related to the density of the stock, which obviously needs to be nonnegative for any strike K.

Example 4.3.1 (Exploiting calendar spread arbitrage) Consider the following European call options, $V_c(t_0, S_0; K, T_1)$ and $V_c(t_0, S_0; K, T_2)$, with $t_0 < T_1 < T_2$, zero interest rate and no dividend payment.

By Jensen's inequality, which states that for any convex function $g(\cdot)$ the following inequality holds

$$g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)],$$

for any random variable X (see also Definition 3.1.1), we will show that $V_c(t_0, S_0; K, T_1) < V_c(t_0, S_0; K, T_2)$.

The value of a call, as seen from time $t = T_1$, with maturity time T_2, is given by:

$$\begin{aligned} V_c(T_1, S(T_1); K, T_2) &= \mathbb{E}^{\mathbb{Q}}\left[\max(S(T_2) - K, 0)\middle|\mathcal{F}(T_1)\right] \\ &\geq \max(\mathbb{E}^{\mathbb{Q}}\left[S(T_2)\middle|\mathcal{F}(T_1)\right] - K, 0). \end{aligned} \tag{4.58}$$

With zero interest rate, $S(t)$ is a martingale and thus $\mathbb{E}^{\mathbb{Q}}\left[S(T_2)\middle|\mathcal{F}(T_1)\right] = S(T_1)$. Therefore, (4.58) gives us:

$$V_c(T_1, S(T_1); K, T_2) \geq \max(S(T_1) - K, 0) = V_c(T_1, S(T_1); K, T_1). \qquad \blacklozenge$$

A call option maturing "at a later time" should thus be more expensive than another one maturing "earlier", which can be generalized to any maturity time T.

Example 4.3.2 Examples for each of the three types of arbitrage are presented in Figure 4.7. The upper left figure illustrates the calendar spread arbitrage. For some strike prices the corresponding call option price with a longer maturity time is cheaper than the option with shorter maturity time. The upper right figure shows a case where the spread arbitrage is present and the lower figure shows a case with butterfly spread arbitrage. ♦

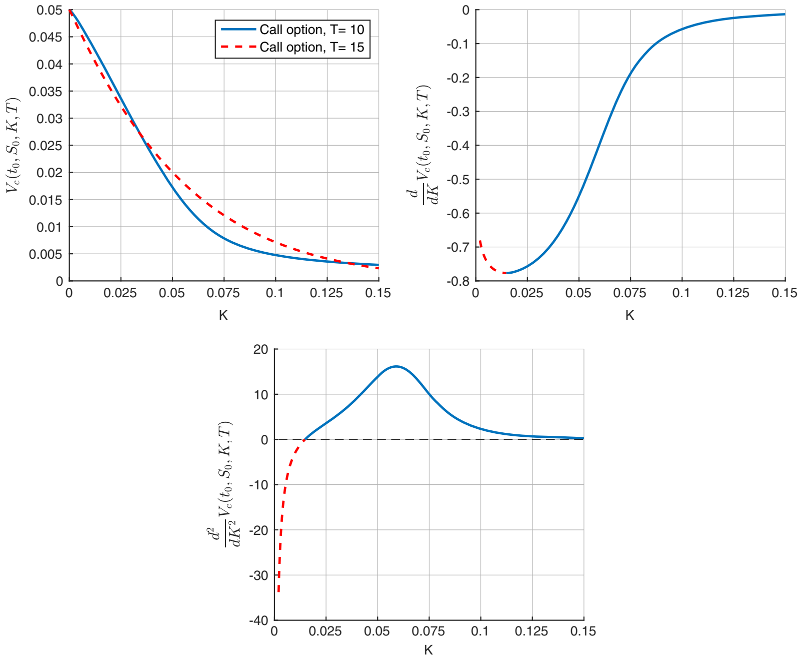


Figure 4.7: *Three types of arbitrage. Upper left: calendar spread arbitrage. Upper right: spread arbitrage. Lower: butterfly spread arbitrage.*

The spread and the butterfly arbitrage can be related to the PDF and CDF of the stock $S(T)$,

$$\frac{\partial}{\partial K}V_c(t_0,S_0;K,T) = \mathrm{e}^{-r(T-t_0)}\left(F_{S(T)}(K)-1\right), \tag{4.59}$$

and,

$$\frac{\partial^2}{\partial K^2}V_c(t_0,S_0;K,T) = \mathrm{e}^{-r(T-t_0)}f_{S(T)}(K). \tag{4.60}$$

As a consequence, by plotting the first and the second derivatives of the call option prices, arbitrage opportunities can immediately be recognized. The CDF of the stock price, $F_{S(T)}(K)$, needs to be monotone and the PDF of the stock price, $f_{S(T)}(K)$, needs to be nonnegative and should integrate to one.

4.3.3 Advanced implied volatility interpolation

The number of implied volatility market quotes for each maturity time T_i will vary for each market and for each asset class. For some stocks there are even hardly any implied volatility market quotes and for some other stocks there are only a few. The number of market quotes depends on the market's liquidity and on the particular market conventions. For example, in foreign-exchange (FX) markets (to be discussed in Chapter 15), one may observe from 3 to 7 implied volatility FX market quotes for each expiry date. The number of market quotes will depend on the particular currency pair. In Table 4.1, an example of FX market implied volatilities is presented. The ATM level (in bold letters) and four quotes for the ITM and OTM options are available.

Now, imagine that a financial institution should give an option price for a strike price which is not quoted in the FX market, e.g. for $K = 145$, which is not present in Table 4.1. For this purpose an interpolation technique is required, to interpolate between the available volatilities. The same holds true when pricing a nonstandard European-style derivative, for which the LV model or the Breeden-Litzenberger method is employed. These two models are based on a *continuum of traded strike prices*, and their implied volatilities.

There are different techniques to interpolate between the available market quotes, like linear, spline, or tangent spline, interpolation. One of the popular methods for the interpolation of implied volatilities is based on a paper by Hagan *et al.* [2002], as this interpolation is built on a specific implied volatility parametrization. This method is popular because of a direct relation between the interpolation and an arbitrage-free SDE model, the Stochastic Alpha Beta Rho (SABR) model.[5]

The parametrization gained in popularity because implied volatilities are analytically available, as an asymptotic expansion formula of an implied volatility was presented. The expression for the Black-Scholes implied volatility approximation

Table 4.1: *Implied volatilities for FX (USD/JPY) for* $T = 1$ *and forward* $S_F(t_0, T) = 131.5$. *The bold letters indicate the ATM level.*

$T = 1y$	$K = 110.0$	$K = 121.3$	$\mathbf{K = 131.5}$	$K = 140.9$	$K = 151.4$
$\sigma_{imp}(T, K)$ [%]	14.2	11.8	**10.3**	10.0	10.7

[5]The SABR model is defined by the following system of SDEs:

$$
\begin{aligned}
\mathrm{d}S_F(t,T) &= \sigma(t)(S_F(t,T))^\beta \mathrm{d}W_F(t), \quad S_F(t_0,T) = S_{F,0},\\
\mathrm{d}\sigma(t) &= \gamma\sigma(t)\mathrm{d}W_\sigma(t), \qquad\quad \sigma(t_0) = \alpha.
\end{aligned}
$$

where $\mathrm{d}W_F(t)\mathrm{d}W_\sigma(t) = \rho\mathrm{d}t$ and the processes are defined under the T-forward measure $\mathbb{Q}^T$. The variance process $\sigma(t)$ is a lognormal distribution. Further, since for a constant $\sigma := \sigma(t)$ the forward $S_F(t,T)$ follows a CEV process, one can expect that the conditional SABR process, $S(T)$ given the paths of $\sigma(t)$ on the interval $0 \leq t \leq T$, is a CEV process as well. Systems of SDEs will be discussed in Chapter 7, and the CEV process in Chapter 14.

of the SABR model[6] reads:

$$\hat{\sigma}(T,K) = \frac{\hat{a}(K)\hat{c}(K)}{g(\hat{c}(K))} \times \left[1 + \left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(S_F(t_0)K)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta\gamma\alpha}{(S_F(t_0)K)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24}\gamma^2\right)T\right], \tag{4.61}$$

where

$$\hat{a}(K) = \frac{\alpha}{(S_F(t_0)\cdot K)^{\frac{1-\beta}{2}}\left(1 + \frac{(1-\beta)^2}{24}\log^2\left(\frac{S_F(t_0)}{K}\right) + \frac{(1-\beta)^4}{1920}\log^4\left(\frac{S_F(t_0)}{K}\right)\right)},$$

and

$$\hat{c}(K) = \frac{\gamma}{\alpha}(S_F(t_0)K)^{\frac{1-\beta}{2}}\log\frac{S_F(t_0)}{K}, \qquad g(x) = \log\left(\frac{\sqrt{1-2\rho x + x^2} + x - \rho}{1-\rho}\right).$$

In the case of *at-the-money options*, i.e. for $S_F(t_0) = K$, the formula reduces to,

$$\hat{\sigma}(T,K) \approx \frac{\alpha}{(S_F(t_0))^{1-\beta}}\left(1 + \left[\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(S_F(t_0))^{2-2\beta}} + \frac{1}{4}\frac{\rho\beta\alpha\gamma}{(S_F(t_0))^{1-\beta}} + \frac{2-3\rho^2}{24}\gamma^2\right]T\right).$$

Because the Hagan formula is derived from a model of an advanced 2D structure (the model falls into the class of the Stochastic Local Volatility models that will be discussed in Chapters 10), it can model a wide range of different implied volatility shapes. In Figure 4.8 the effect of different model parameters on the implied volatility shapes is shown. Notice that both parameters β and ρ have an effect on the implied volatility skew. In practice, β is often fixed, whereas ρ is used in a calibration. Parameter α controls the level of the implied volatility smile and γ the magnitude of the curvature of the smile.

Example 4.3.3 (Interpolation of implied volatilities) In this example, we will use the market implied volatility quotes from Table 4.1. For the given set of quotes, we compare two interpolation techniques for the implied volatilities, the linear interpolation with the Hagan interpolation. The linear interpolation is straightforward as it directly interpolates between any two market implied volatilities. For the Hagan interpolation the procedure is more involved. First, the model parameters $\alpha, \beta, \rho, \gamma$ need to be determined for which the parametrization fits best to the market quotes, i.e., minimize the difference $|\hat{\sigma}(T,K_i) - \sigma_{imp}(T,K_i)|$, for all K_i.

[6]Note that this is the *Black-Scholes* implied volatility and *not* the SABR implied volatility.

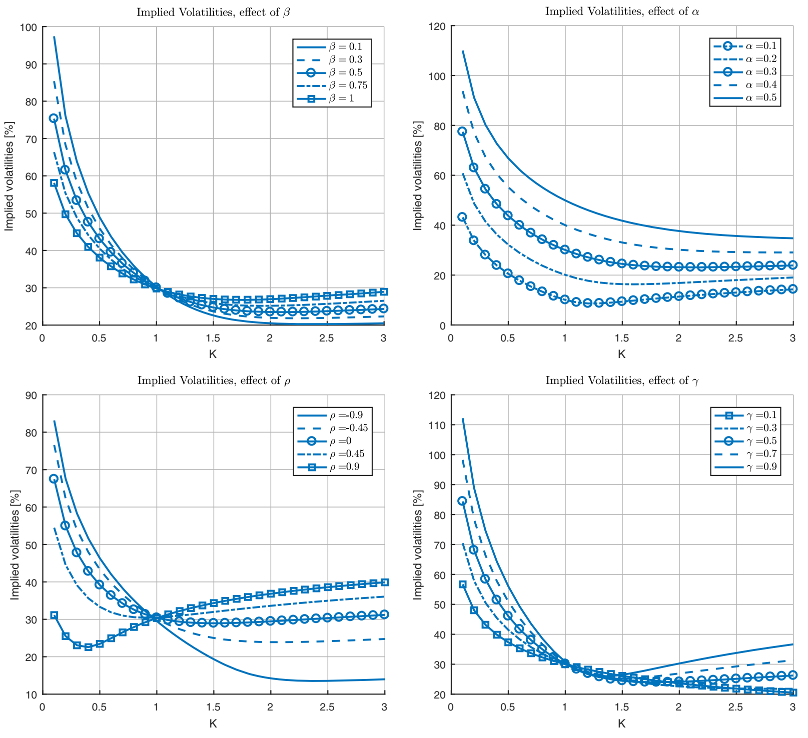


Figure 4.8: *Different implied volatility shapes under Hagan's implied volatility parametrization, depending on different model parameters.*

In Figure 4.9 the results of the two interpolation methods are presented. The left side figure shows the market implied volatilities, with the linear interpolation and the Hagan interpolation. Both interpolation curves seem to fit very well to the market quotes. We also discussed that when market call and put option prices are available, it is easy to approximate the underlying density function, based on these quotes, as

$$f_{S(T)}(K) = \mathrm{e}^{r(T-t_0)} \frac{\partial^2 V_c(t_0, S_0; K, T)}{\partial K^2}.$$

The right side figure shows the corresponding *"implied densities"* from the two interpolation techniques. Notice the nonsmooth nature of the density which is

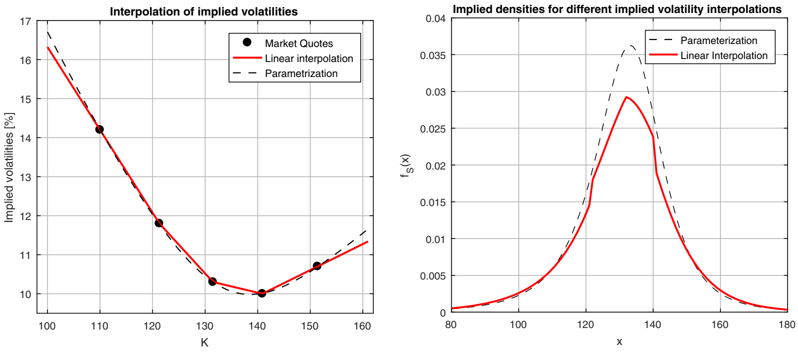


Figure 4.9: *Different interpolation techniques applied to FX market quotes. Left: interpolated implied volatilities. Right: the corresponding implied densities.*

based on linear interpolation. Integration of these density functions yields,

$$\text{Hagan}: \int_{\mathbb{R}} f_{S(T)}(z)\mathrm{d}z = 0.9969, \quad \text{Linear}: \int_{\mathbb{R}} f_{S(T)}(z)\mathrm{d}z = 0.8701.$$

In fact, the linear interpolation gives rise to a *"butterfly spread arbitrage"*, and is therefore unsatisfactory. The Hagan interpolation technique, on the other hand, yields much better results. However, this interpolation is also not perfect and it may give unsatisfactory results, in extreme market settings (for which, due to the parameter settings, the asymptotic volatility formula is not accurate anymore), like for example in the case of negative interest rates. The interpolation routine can however be improved, as discussed in [Grzelak and Oosterlee, 2016].

The SABR parametrization is often preferred to a standard linear or spline interpolation. Because the parametrization is derived from an arbitrage-free model, it fits to the purpose of interpolating market implied volatilities.

When interpolating implied volatilities, it is important that *interpolation arbitrage* is avoided. An implied volatility parametrization which is derived from an arbitrage-free asset price model is often preferred to a standard interpolation technique which is available in the numerical software packages.

♦

4.3.4 Simulation of local volatility model

The expression for the local volatility is used within the SDE,

$$\boxed{\mathrm{d}S(t) = rS(t)\mathrm{d}t + \sigma_{LV}(t, S(t))S(t)\mathrm{d}W(t).}$$

So, the local volatility term (4.49) at each time step $T = t$ in $K = S(t)$ is to be determined.

The local volatility framework relies on the available implied volatility surface, $\sigma_{\text{imp}}(T, K)$, for each expiry date T and strike price K. For each underlying only a few option market quotes are available, so that the industry standard is to parameterize the implied volatility surface by a 2D continuous function $\hat{\sigma}(T, K)$, which is calibrated to the market quotes, $\sigma_{\text{imp}}(T_i, K_j)$, at the set of available expiry dates and strike prices. The volatility parametrization $\hat{\sigma}(T, K)$ needs to satisfy the arbitrage-free conditions.

Let us consider the Hagan interpolation (4.61), using $\rho = 0, \beta = 1$, for which the implied volatilities are given in the following form:

$$\hat{\sigma}(T, K) = \frac{\gamma \log(S_F(t_0)/K) \left(1 + \frac{\gamma^2}{12} T\right)}{\log\left(\sqrt{1 + (\frac{\gamma}{\alpha} \log(S_F(t_0)/K))^2} + \frac{\gamma}{\alpha} \log(S_F(t_0)/K)\right)}, \tag{4.62}$$

and

$$\hat{\sigma}(T, S_F(t_0)) = \alpha \left(1 + \frac{\gamma^2}{12} T\right),$$

This functional form in (4.62) is commonly used in the industry [Grzelak and Oosterlee, 2016].

We start with the simulation of the local volatility model from (4.40), with $\sigma_{LV}(t, S)$ as defined in (4.56), and where $\sigma_{\text{imp}}(t, S(t)) = \hat{\sigma}(t, S(t))$ in (4.62).

The simulation of Equation (4.40) can be performed by means of an Euler discretization, which will be discussed in more detail in Chapter 9, i.e.,

$$s_{i+1} = s_i + r s_i \Delta t + \sigma_{LV}(t_i, s_i) s_i \big(W(t_{i+1}) - W(t_i)\big), \tag{4.63}$$

with $\sigma_{LV}(t_i, s_i)$ given by (4.56), i.e.

$$\sigma_{LV}^2(t_i, s_i) = \frac{\dfrac{\partial w}{\partial t_i} + r s_i \dfrac{\partial w}{\partial s_i}}{1 + s_i \dfrac{\partial w}{\partial s_i}\left(\dfrac{1}{2} - \dfrac{y}{w}\right) + \dfrac{1}{2} s_i^2 \dfrac{\partial^2 w}{\partial s_i^2} + \dfrac{1}{2} s_i^2 \left(\dfrac{\partial w}{\partial s_i}\right)^2 \left(-\dfrac{1}{8} - \dfrac{1}{2w} + \dfrac{y^2}{2w^2}\right)}, \tag{4.64}$$

for $t_0 = 0$ and $w := t_i \hat{\sigma}^2(t_i, s_i)$,

$$\frac{\partial w}{\partial t_i} = \hat{\sigma}(t_i, s_i)^2 + 2t\hat{\sigma}(t_i, s_i) \frac{\partial \hat{\sigma}(t_i, s_i)}{\partial t_i}, \quad \frac{\partial w}{\partial s_i} = 2t_i \hat{\sigma}(t_i, s_i) \frac{\partial \hat{\sigma}(t_i, s_i)}{\partial s_i},$$

$$\frac{\partial^2 w}{\partial s_i^2} = 2t_i \left(\frac{\partial \hat{\sigma}(t_i, s_i)}{\partial s_i}\right)^2 + 2t_i \hat{\sigma}(t_i, s_i) \frac{\partial^2 \hat{\sigma}(t_i, s_i)}{\partial s_i^2}.$$

The local volatility component $\sigma_{LV}(t_i, s_i)$ in (4.64) needs to be evaluated for each realization s_i. In the case of a Monte Carlo simulation when many thousands asset paths are generated, the evaluation of (4.64) can become computationally intensive.

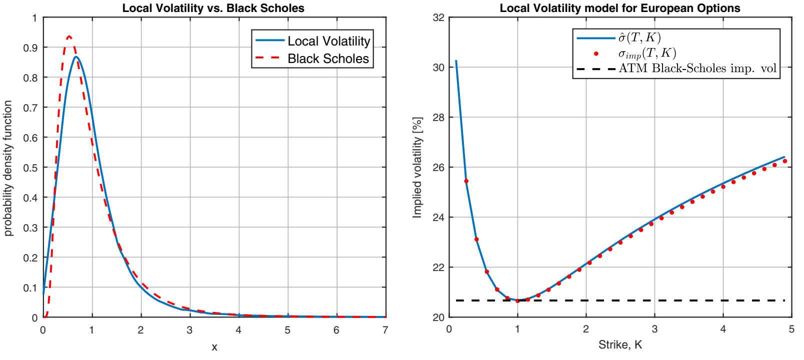


Figure 4.10: *Local volatility versus Black Scholes model. Left: the stock densities of the Black-Scholes and the local volatility model. Right: volatilities $\hat{\sigma}(T,K)$ and $\sigma_{imp}(T,K)$ obtained by the local volatility model, with $T = 10$, $\gamma = 0.2$, $S_0 = 1$ and $\alpha = 0.2$.*

As a sanity check, when the model has been implemented correctly the input volatilities $\hat{\sigma}(T,K)$ should resemble $\sigma_{\text{imp}}(T,K)$.

In the left side picture in Figure 4.10, we compare the local volatility density with the Black-Scholes log-normal density function in which the ATM volatility is substituted. In the right side figure the input volatilities $\hat{\sigma}(T,K)$ and the output implied volatilities $\sigma_{\text{imp}}(T,K)$, as obtained from the Monte Carlo simulation are shown. The local volatility model generates implied volatilities that resemble the input in the parametrization of $\sigma_{LV}(T,K)$ very well.

We conclude with the following general statements regarding the LV model. Local volatility models resemble very well the implied volatilities that we observe in market option quotes, and they are used as an input to the LV model. Calibration to these input European options is thus highly accurate, and the model is “tailored” to these available option quotes. However, LV models may suffer from significant mispricing inaccuracy when dealing with financial derivatives products that depend on the volatility paths and, generally, on transition density functions. This information is simply not “coded” in the LV model framework. In Chapter 10 this problem will be discussed in further detail.

4.4 Exercise set

Exercise 4.1 Employ the Newton-Raphson iteration, the combined root-finding method, as well as Brent's method to find the roots of the following problems,

a. Compute the two solutions of the nonlinear equation,

$$g(x) = \frac{e^x + e^{-x}}{2} - 2x = 0.$$

Present the respective approximations for each iteration (so, the convergence) to the two solutions.

b. Acquire some option market data, either electronically or via a newspaper, of a company whose name starts with your initial or the first letter of your surname. Confirm that the option data contain an implied volatility skew or smile. Plot and discuss the implied volatility as a function of the strike price.

Exercise 4.2 Choose a set of option parameters for a deep in the money option. Confirm that the Newton-Raphson iteration has convergence issues, whereas the combined root-finding method converges in a robust way.

Exercise 4.3 With two new independent variables,

$$X_F := \log\left(\frac{Se^{r(T-t)}}{K}\right), \text{ and } t_* := \sigma\sqrt{T-t},$$

a. Rewrite the arguments $d_{1,2}$ in the Black-Scholes solution, in two variables that only depend on X_F and t_*.

b. Give an interpretation of the inequality $X_F \leq 0$.

c. Show that for the scaled put option value, $p = V_p(t, S(t))/S(t)$, we have,

$$p(X_F, t_*) = e^{-X_F} F_{\mathcal{N}(0,1)}(-d_2) - F_{\mathcal{N}(0,1)}(-d_1).$$

d. Determine

$$\frac{\partial p(X_F, t_*)}{\partial t_*}$$

and give an interpretation of this derivative.

Exercise 4.4 For $c(y, \omega) := S_0\Big[F_{\mathcal{N}(0,1)}(d_1) - \mathrm{e}^y F_{\mathcal{N}(0,1)}(d_2)\Big]$, with

$$d_1 = -\frac{y}{\sqrt{\omega}} + \frac{1}{2}\sqrt{\omega},$$

and $d_2 = d_1 - \sqrt{\omega}$, and y and ω defined in (4.50) show the following equality:

$$\frac{\partial^2 c}{\partial w^2} = \frac{\partial c}{\partial w}\left(-\frac{1}{8} - \frac{1}{2w} + \frac{y^2}{2w^2}\right).$$

Exercise 4.5 Use the results from Exercise 4.4 to show that:

$$\frac{\partial^2 c}{\partial w \partial y} = \frac{\partial c}{\partial w}\left(\frac{1}{2} - \frac{y}{w}\right), \quad \text{and} \quad \frac{\partial^2 c}{\partial y^2} = \frac{\partial c}{\partial y} + 2\frac{\partial c}{\partial w}.$$

Exercise 4.6 For which of the following payoff functions can option pricing be done in the Breeden-Litzenberger framework?

a. A digital option, where the payoff function is based on $\mathbb{1}_{S(T)>K}$,

b. A spread option based on the performance of two stocks, where the payoff reads $\max(S_1(T) - S_2(T), K)$,

c. A performance option, with as the payoff, $\max(S(T_2)/S(T_1) - K, 0)$,

d. A cliquet-type option, with payoff function, $\min(\max(\frac{S(T)}{S(t_0)} - K_1, 0), K_2)$.

Exercise 4.7 Let us assume a market in which $r = 0$, $S(t_0) = 1$ and the implied volatilities for $T = 4$ are given by the following formula,

$$\hat{\sigma}(K) = 0.510 - 0.591K + 0.376K^2 - 0.105K^3 + 0.011K^4, \tag{4.65}$$

with an upper limit, which is given by $\hat{\sigma}(K) = \hat{\sigma}(3)$ for $K > 3$.

In the Breeden-Litzenberger framework any European-type payoff function can be priced with the following formula, in the case $r = 0$,

$$V(t_0, S_0) = V(T, S_0) + \int_0^{S_0} V_p(t_0, x)\frac{\partial^2 V(T,x)}{\partial x^2}\mathrm{d}x + \int_{S_0}^{\infty} V_c(t_0, x)\frac{\partial^2 V(T,x)}{\partial x^2}\mathrm{d}x, \tag{4.66}$$

where $V(T, x)$ is one of the payoff functions at time T, that are given below. $V_p(t_0, x)$ and $V_c(t_0, x)$ are the put and call prices with strike price x. These prices can be obtained by evaluating the Black-Scholes formula for the volatilities in Equation (4.65).

Using Equation (4.66) compute numerically the option prices at $t = 0$, for the following payoff functions,

a. $V(T, S(T)) = \max(S^2(T) - 1.2S(T), 0)$

b. $V(T, S(T)) = \max(S(T) - 1.5, 0)$

c. $V(T, S(T)) = \max(1.7 - S(T), 0) + \max(S(T) - 1.4, 0)$

d. $V(T, S(T)) = \max(4 - S^3(T), 0) + \max(S(T) - 2, 0)$

Hint: Approximate the derivatives of the payoff $V(T, x)$, by finite differences. Note that increment δ, in $V(T, x + \delta)$, cannot be too small for accuracy reasons.

For which of the payoff functions above can the option price at $t = 0$ be determined without using the Breeden-Litzenberger model?

Exercise 4.8 Consider the SABR formula for the Black-Scholes implied volatility, i.e., Equation (4.61) with the following set of parameters, using $S_F(t_0) = S(t_0)\mathrm{e}^{r(T-t_0)}$, $T = 2.5$, $\beta = 0.5$, $S(t_0) = 5.6$, $r = 0.015$, $\alpha = 0.2$, $\rho = -0.7$, $\gamma = 0.35$. Based on the relation between option prices and the stock's probability density function,

$$\begin{aligned} f_{S(T)}(x) &= \mathrm{e}^{r(T-t_0)}\frac{\partial^2 V_c(t_0, S_0; x, T)}{\partial x^2} \\ &\approx \mathrm{e}^{r(T-t_0)}\frac{V_c(t_0, S_0; x+\Delta x, T) - 2V_c(t_0, S_0; x, T) + V_c(t_0, S_0; x-\Delta x, T)}{\Delta x^2}, \end{aligned} \tag{4.67}$$

compute the probability density function $f_{S(T)}(x)$. Is this probability density function free of arbitrage? Explain the impact of Δx on the quality of the PDF.

Exercise 4.9 With the parameters, $T = 1$, $S_0 = 10.5$, $r = 0.04$, we observe the following set of implied volatilities in the market,

Table 4.2: *Implied volatilities and the corresponding strikes.*

K	3.28	5.46	8.2	10.93	13.66	16.39	19.12	21.86
$\sigma_{imp}(T,K)[\%]$	31.37	22.49	14.91	9.09	6.85	8.09	9.45	10.63

Consider the following interpolation routines,

a) linear,

b) cubic spline,

c) nearest neighbor.

Which of these interpolations will give rise to the smallest values for the butterfly arbitrage? Explain how to reduce the arbitrage values. *Hint:* Assume a flat extrapolation prior to the first and behind the last strike price.

Exercise 4.10 Consider two expiry dates, $T_1 = 1y$ and $T_2 = 2y$, with as the parameter values, $r = 0.1$, $S_0 = 10.93$ and the following implied volatilities,

Table 4.3: *Implied volatilities and the corresponding strikes.*

K	3.28	5.46	8.2	10.93	13.66	16.39	19.12	21.86
$\sigma_{imp}(T_1,K)[\%]$	31.37	22.49	14.91	9.09	6.85	8.09	9.45	10.63
$\sigma_{imp}(T_2,K)[\%]$	15.68	11.25	7.45	4.54	3.42	4.04	4.72	5.31

Check for a calendar arbitrage based on these values.

Exercise 4.11 Consider two independent stock prices, $S_1(t)$ and $S_2(t)$, determine the following expectations using the Breeden-Litzenberger model,

$$\mathbb{E}^{\mathbb{Q}}\left[\log\left(S_1(T)^{S_2(T)}\right)\Big|\mathcal{F}(t_0)\right], \quad \mathbb{E}^{\mathbb{Q}}\left[\log\prod_{i=1}^{N}\frac{S_1(t_{i+1})}{S_1(t_i)}\frac{S_2(t_{i+1})}{S_2(t_i)}\Big|\mathcal{F}(t_0)\right].$$

Exercise 4.12 When pricing financial options based on the market probability density function, it is possible to use either Equation (4.11) or Equation (4.20). With the following payoff function, $V(T,S) = \max(S^2 - K, 0)$, and the settings as in Exercise 4.9, give arguments which of the two representations is more stable numerically, and present some corresponding numerical results.

This page intentionally left blank

CHAPTER 5

Jump Processes

In this chapter:

One way to deal with the market observed *implied volatility smile* is to consider asset prices based on jump processes. We first explain the *jump diffusion processes* in **Section 5.1** and detail the generalization of Itô's lemma to jump processes. The option valuation equations under these asset price processes with jumps turn out to be *partial integro-differential equations*. An analytic solution can be derived, based on the Feynman-Kac Theorem (see **Section 5.2**). The jump diffusion models can be placed in the class of exponential Lévy asset price processes. They are discussed in **Section 5.3**. We discuss Lévy jump models with a *finite* (**Section 5.3.1**) or an *infinite number of jumps* (**Section 5.4**). For certain specific Lévy asset models, numerical examples give insight in the impact of the different model parameters on the asset price density and the implied volatility. We end this chapter, in Section 5.5, with a general discussion on the use of jumps for asset price modeling in **Section 5.5**.
Important is that for the asset models in this chapter, the *characteristic function* can be derived. The characteristic function will form the basis of highly efficient Fourier option pricing techniques, in Chapter 6.

Keywords: alternative model for asset prices, jump diffusion process, Poisson process, Lévy jump models, impact on option pricing PDE, characteristic function, Itô's lemma.

5.1 Jump diffusion processes

We analyze the Black-Scholes model extended by independent jumps, that are driven by a Poisson process, the so-called *jump diffusion process*, and consider the following dynamics for the log-stock process, $X(t) = \log S(t)$, under the real-world measure $\mathbb{P}$:

$$\boxed{\mathrm{d}X(t) = \mu \mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{P}}(t) + J\mathrm{d}X_{\mathcal{P}}(t),} \tag{5.1}$$

where μ is the drift, σ is the volatility, $X_{\mathcal{P}}(t)$ is a Poisson process and variable J gives the jump magnitude; J is governed by a distribution, F_J, of magnitudes. Processes $W^{\mathbb{P}}(t)$ and $X_{\mathcal{P}}(t)$ are assumed to be *independent*.

Before we start working with these dynamics, we first need to discuss some details of jump processes, in particular of the Poisson process, and also the relevant version of Itô's lemma.

Definition 5.1.1 (Poisson random variable) *A Poisson random variable, which is denoted by $X_{\mathcal{P}}$, counts the number of occurrences of an event during a given time period. The probability of observing $k \geq 0$ occurrences in the time period is given by*

$$\mathbb{P}[X_{\mathcal{P}} = k] = \frac{\xi_p^k \mathrm{e}^{-\xi_p}}{k!}.$$

The mean of $X_{\mathcal{P}}$, $\mathbb{E}[X_{\mathcal{P}}] = \xi_p$, gives the average number of occurrences of the event, while for the variance, $\mathbb{V}\mathrm{ar}[X_{\mathcal{P}}] = \xi_p$. ◀

Definition 5.1.2 (Poisson process) *A Poisson process, $\{X_{\mathcal{P}}(t), t \geq t_0 = 0\}$, with parameter $\xi_p > 0$ is an integer-valued stochastic process, with the following properties*

- $X_{\mathcal{P}}(0) = 0$;
- $\forall\ t_0 = 0 < t_1 < \ldots < t_n$, *the increments* $X_{\mathcal{P}}(t_1) - X_{\mathcal{P}}(t_0), X_{\mathcal{P}}(t_2) - X_{\mathcal{P}}(t_1), \ldots, X_{\mathcal{P}}(t_n) - X_{\mathcal{P}}(t_{n-1})$ *are independent random variables;*
- *for $s \geq 0, t > 0$ and integers $k \geq 0$, the increments have the Poisson distribution:*

$$\mathbb{P}[X_{\mathcal{P}}(s+t) - X_{\mathcal{P}}(s) = k] = \frac{(\xi_p t)^k \mathrm{e}^{-\xi_p t}}{k!}. \tag{5.2}$$

◀

The Poisson process $X_{\mathcal{P}}(t)$ is a counting process, with the number of *jumps* in any time period of length t specified via (5.2). Equation (5.2) confirms the stationary increments, since the increments only depend on the length of the interval and not on initial time s.

Parameter ξ_p is the *rate* of the Poisson process, i.e. it indicates the number of jumps in a time period. The probability that exactly one event occurs in a small time interval, $\mathrm{d}t$, follows from (5.2) as

$$\mathbb{P}[X_{\mathcal{P}}(s+\mathrm{d}t) - X_{\mathcal{P}}(s) = 1] = \frac{(\xi_p \mathrm{d}t)\mathrm{e}^{-\xi_p \mathrm{d}t}}{1!} = \xi_p \mathrm{d}t + o(\mathrm{d}t),$$

and the probability that no event occurs in $\mathrm{d}t$ is

$$\mathbb{P}[X_\mathcal{P}(s+\mathrm{d}t) - X_\mathcal{P}(s) = 0] = \mathrm{e}^{-\xi_p \mathrm{d}t} = 1 - \xi_p \mathrm{d}t + o(\mathrm{d}t).$$

In a time interval $\mathrm{d}t$, a jump will arrive with probability $\xi_p \mathrm{d}t$, resulting in:

$$\mathbb{E}[\mathrm{d}X_\mathcal{P}(t)] = 1 \cdot \xi_p \mathrm{d}t + 0 \cdot (1 - \xi_p \mathrm{d}t) = \xi_p \mathrm{d}t,$$

where $\mathrm{d}X_\mathcal{P}(t) = X_\mathcal{P}(s+\mathrm{d}t) - X_\mathcal{P}(s)$. The expectation is thus given by

$$\mathbb{E}[X_\mathcal{P}(s+t) - X_\mathcal{P}(s)] = \xi_p t.$$

With $X_\mathcal{P}(0) = 0$, the expected number of events in a time interval with length t, setting $s = 0$, equals

$$\mathbb{E}[X_\mathcal{P}(t)] = \xi_p t. \tag{5.3}$$

If we define another process, $\bar{X}_\mathcal{P}(t) := X_\mathcal{P}(t) - \xi_p t$, then $\mathbb{E}[\mathrm{d}\bar{X}_\mathcal{P}(t)] = 0$, so that process $\bar{X}_\mathcal{P}(t)$, which is referred to as the *compensated* Poisson process, is a martingale.

Example 5.1.1 (Paths of the Poisson process) We present some discrete paths that have been generated by a Poisson process, with $\xi_p = 1$. The left-hand picture in Figure 5.1 displays the paths by $\mathrm{d}X_\mathcal{P}(t)$, whereas the right-hand picture shows the same paths for the compensated Poisson process, $-\xi_p \mathrm{d}t + \mathrm{d}X_\mathcal{P}(t)$.

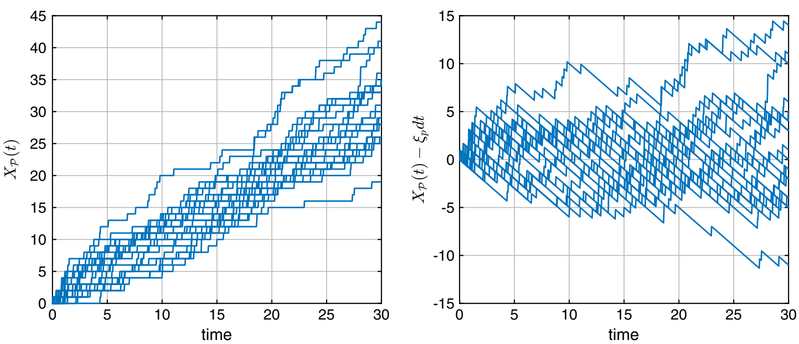


Figure 5.1: *Monte Carlo paths for the Poisson (left) and the compensated Poisson process (right), $\xi_p = 1$.*

Given the following SDE:

$$\mathrm{d}X(t) = J(t)\mathrm{d}X_{\mathcal{P}}(t), \tag{5.4}$$

we may define the stochastic integral with respect to the Poisson process $X_{\mathcal{P}}(t)$, by

$$X(T) - X(t_0) = \int_{t_0}^{T} J(t)\mathrm{d}X_{\mathcal{P}}(t) := \sum_{k=1}^{X_{\mathcal{P}}(T)} J_k. \tag{5.5}$$

Variable J_k for $k \geq 1$ is an i.i.d. sequence of random variables with a jump-size probability distribution F_J, so that $\mathbb{E}[J_k] = \mu_J < \infty$.

5.1.1 Itô's lemma and jumps

In Chapter 2, we have discussed that for traded assets the martingale property should be satisfied for the discounted asset price $S(t)/M(t)$, under the risk neutral pricing measure, see Section 2.3. Just substituting $\mu = r$ in (5.1) is however not sufficient to achieve this, as the jumps have an impact on the drift term too. A *drift adjustment* needs to compensate for average jump sizes.

In order to derive the dynamics for $S(t) = \exp(X(t))$, with $X(t)$ in (5.1), a variant of Itô's lemma which is related to the Poisson process, needs to be employed. We present a general result below for models with independent Poisson jumps.

Result 5.1.1 (Itô's lemma for Poisson process) *We consider a cádlág process, $X(t)$, defined as:*

$$\mathrm{d}X(t) = \bar{\mu}(t, X(t))\mathrm{d}t + \bar{J}(t, X(t_-))\mathrm{d}X_{\mathcal{P}}(t), \quad \textit{with} \quad X(t_0) \in \mathbb{R}, \tag{5.6}$$

where $\bar{\mu}, \bar{J} : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ are deterministic, continuous functions and $X_{\mathcal{P}}(t)$ is a Poisson process, starting at $t_0 = 0$.
For a differentiable function, $g : [0,\infty) \times \mathbb{R} \to \mathbb{R}$, the Itô differential is given by:

$$\begin{aligned}\mathrm{d}g(t, X(t)) = {} & \left[\frac{\partial g(t, X(t))}{\partial t} + \bar{\mu}(t, X(t))\frac{\partial g(t, X(t))}{\partial X}\right]\mathrm{d}t \\ & + \Big[g\left(t, X(t_-) + \bar{J}(t, X(t_-))\right) - g(t, X(t_-))\Big]\mathrm{d}X_{\mathcal{P}}(t),\end{aligned} \tag{5.7}$$

where the left limit is denoted by $X(t_-) := \lim_{s\to t} X(s), s < t$, so that, by the continuity of $\bar{J}(\cdot)$, its left limit equals $\bar{J}(t, X(t_-))$.

An intuitive explanation for Itô's formula in the case of jumps is that when a jump takes place, i.e. $\mathrm{d}X_{\mathcal{P}}(t) = 1$, the process "jumps" from $X(t_-)$ to $X(t)$, with the jump size determined by function $\bar{J}(t, X(t))$ [Sennewald and Wälde, 2006], resulting in the following relation:

$$g(t, X(t)) = g\left(t, X(t_-) + \bar{J}(t, X(t_-)\right).$$

After the jump at time t, the function $g(\cdot)$ is adjusted with the jump size, which was determined at time t_-.

In practical applications, stochastic processes that include both a Brownian motion and a Poisson process are encountered, as in Equation (5.1). The general dynamics of this combined process are given by:

$$\mathrm{d}X(t) = \bar{\mu}(t, X(t))\mathrm{d}t + \bar{J}(t, X(t_-))\mathrm{d}X_{\mathcal{P}}(t) + \bar{\sigma}(t, X(t))\mathrm{d}W(t), \quad \text{with} \quad X(t_0) \in \mathbb{R}, \tag{5.8}$$

with $\bar{\mu}(\cdot,\cdot)$ the drift, $\bar{\sigma}(\cdot,\cdot)$ the diffusion term, and $\bar{J}(\cdot,\cdot)$ the jump magnitude function. For a function $g(t, X(t))$ in (5.7) an extension of Result 5.1.1 is then required. Assuming that the Poisson process $X_{\mathcal{P}}(t)$ is *independent* of the Brownian motion $W(t)$, the dynamics of $g(t, X(t))$ are given by:

$$\boxed{\begin{aligned}\mathrm{d}g(t, X(t)) = &\left[\frac{\partial g(t, X(t))}{\partial t} + \bar{\mu}(t, X(t))\frac{\partial g(t, X(t))}{\partial X}\right. \\ &+ \left.\frac{1}{2}\bar{\sigma}^2(t, X(t))\frac{\partial^2 g(t, X(t))}{\partial X^2}\right]\mathrm{d}t \\ &+ \left[g\big(t, X(t_-) + \bar{J}(t, X(t_-))\big) - g(t, X(t_-))\right]\mathrm{d}X_{\mathcal{P}}(t) \\ &+ \bar{\sigma}(t, X(t))\frac{\partial g(t, X(t))}{\partial X}\mathrm{d}W(t), \qquad (5.9)\end{aligned}}$$

Here we made use of the Itô multiplication Table 5.1, where the cross terms involving the Poisson process are also handled. An intuitive way to understand the Poisson process rule in Table 5.1 is found in the notion that the term $\mathrm{d}X_{\mathcal{P}} = 1$ with probability $\xi_p\mathrm{d}t$, and $\mathrm{d}X_{\mathcal{P}} = 0$ with probability $(1-\xi_p\mathrm{d}t)$, which implies that

$$\begin{aligned}(\mathrm{d}X_{\mathcal{P}})^2 &= \begin{cases} 1^2 & \text{with probability } \xi_p\mathrm{d}t, \\ 0^2 & \text{with probability } (1-\xi_p\mathrm{d}t) \end{cases} \\ &= \mathrm{d}X_{\mathcal{P}}.\end{aligned}$$

To apply Itô's lemma to the function $S(t) = \mathrm{e}^{X(t)}$, with $X(t)$ in (5.1), we substitute $\bar{\mu}(t, X(t)) = \mu$, $\bar{\sigma}(t, X(t)) = \sigma$ and $\bar{J}(t, X(t_-)) = J$ in (5.9), giving,

$$\mathrm{d}\mathrm{e}^{X(t)} = \left(\mu\mathrm{e}^{X(t)} + \frac{1}{2}\sigma^2\mathrm{e}^{X(t)}\right)\mathrm{d}t + \sigma\mathrm{e}^{X(t)}\mathrm{d}W(t) + \left(\mathrm{e}^{X(t)+J} - \mathrm{e}^{X(t)}\right)\mathrm{d}X_{\mathcal{P}}(t),$$

Table 5.1: *Itô multiplication table for Poisson process.*

	$\mathrm{d}t$	$\mathrm{d}W(t)$	$\mathrm{d}X_{\mathcal{P}}(t)$
$\mathrm{d}t$	0	0	0
$\mathrm{d}W(t)$	0	$\mathrm{d}t$	0
$\mathrm{d}X_{\mathcal{P}}(t)$	0	0	$\mathrm{d}X_{\mathcal{P}}(t)$

so that we obtain:

$$\frac{\mathrm{d}S(t)}{S(t)} = \left(\mu + \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}W(t) + \left(\mathrm{e}^J - 1\right)\mathrm{d}X_{\mathcal{P}}(t).$$

Until now, we have derived the dynamics for the stock $S(t)$ under the real-world measure $\mathbb{P}$. The next step is to derive the dynamics of the stock under the risk-neutral measure $\mathbb{Q}$.

For this, we check under which conditions the process $Y(t) := S(t)/M(t)$ is a martingale, or, in other words, the dynamics $dY(t) = \frac{dS(t)}{M(t)} - \frac{rS(t)dt}{M(t)}$, with $\mathrm{d}Y(t) = Y(t+\mathrm{d}t) - Y(t)$, should have an expected value equal to zero:

$$\begin{aligned}\mathbb{E}\left[\mathrm{d}Y(t)\right] = \mathbb{E}\left[\mu S(t) + \frac{1}{2}\sigma^2 S(t) - rS(t)\right]\mathrm{d}t + \mathbb{E}\left[\sigma S(t)\mathrm{d}W(t)\right]\\ + \mathbb{E}\left[\left(\mathrm{e}^J - 1\right)S(t)\mathrm{d}X_{\mathcal{P}}(t)\right].\end{aligned}$$

From the properties of the Brownian motion and Poisson process, and the fact that all random components in the expression above are mutually independent,[1] we get:

$$\begin{aligned}\mathbb{E}\left[\mathrm{d}Y(t)\right] &= \mathbb{E}\left[\mu S(t) + \frac{1}{2}\sigma^2 S(t) - rS(t)\right]\mathrm{d}t + \mathbb{E}\left[\left(\mathrm{e}^J - 1\right)S(t)\right]\xi_p\mathrm{d}t\\ &= \left(\mu - r + \frac{1}{2}\sigma^2 + \mathbb{E}\left[\xi_p(\mathrm{e}^J - 1)\right]\right)\mathbb{E}[S(t)]\mathrm{d}t.\end{aligned}$$

By substituting $\mu = r - \frac{1}{2}\sigma^2 - \xi_p\mathbb{E}\left[\mathrm{e}^J - 1\right]$, we have $\mathbb{E}[\mathrm{d}Y(t)] = 0$.

The term $\bar{\omega} := \xi_p\mathbb{E}\left[\mathrm{e}^J - 1\right]$ is the so-called *drift correction term*, which makes the process a martingale.

The dynamics for stock $S(t)$ under the risk-neutral measure $\mathbb{Q}$ are therefore given by:

$$\boxed{\frac{\mathrm{d}S(t)}{S(t)} = \left(r - \xi_p\mathbb{E}\left[\mathrm{e}^J - 1\right]\right)\mathrm{d}t + \sigma\mathrm{d}W^{\mathbb{Q}}(t) + \left(\mathrm{e}^J - 1\right)\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t).} \tag{5.10}$$

The process in (5.10) is often presented in the literature as the *standard jump diffusion model*.

The standard jump diffusion model is directly connected to the following $\mathrm{d}X(t)$ dynamics:

$$\boxed{\mathrm{d}X(t) = \left(r - \xi_p\mathbb{E}\left[\mathrm{e}^J - 1\right] - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}W^{\mathbb{Q}}(t) + J\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t). \quad (5.11)}$$

[1] It can be shown that if $W(t)$ is a Brownian motion and $X_{\mathcal{P}}(t)$ a Poisson process with intensity ξ_p, and both processes are defined on the same probability space $(\Omega, \mathcal{F}(t), \mathbb{P})$, then the processes $W(t)$ and $X_{\mathcal{P}}(t)$ are independent.

5.1.2 PIDE derivation for jump diffusion process

We will determine the pricing PDE in the case the underlying dynamics are driven by the jump diffusion process. We depart from the following SDE:

$$\mathrm{d}S(t) = \bar{\mu}(t, S(t))\mathrm{d}t + \bar{\sigma}(t, S(t))\mathrm{d}W^{\mathbb{Q}}(t) + \bar{J}(t, S(t))\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t), \tag{5.12}$$

where, in the context of Equation (5.10), the functions $\bar{\mu}(t, S(t))$, $\bar{J}(t, S(t))$ and $\bar{\sigma}(t, S(t))$ are equal to

$$\begin{aligned} &\bar{\mu}(t, S(t)) := \left(r - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right]\right) S(t), \quad \bar{\sigma}(t, S(t)) := \sigma S(t), \\ &\bar{J}(t, S(t)) := (\mathrm{e}^J - 1)S(t). \end{aligned} \tag{5.13}$$

The dynamics of process (5.12) are under the risk-neutral measure $\mathbb{Q}$, so that we may apply the martingale approach to derive the option pricing equation under the jump diffusion asset price dynamics. This means that, for a certain payoff $V(T, S)$, the following equality has to hold:

$$\frac{V(t, S)}{M(t)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V(T, S)}{M(T)}\Big|\mathcal{F}(t)\right]. \tag{5.14}$$

$V(t, S)/M(t)$ in (5.14) can be recognized as a martingale under the $\mathbb{Q}$-measure, so that the governing dynamics should not contain any drift terms. With Itô's lemma, the dynamics of V/M, using $M \equiv M(t), V \equiv V(t, S)$, are given by

$$\mathrm{d}\frac{V}{M} = \frac{1}{M}\mathrm{d}V - r\frac{V}{M}\mathrm{d}t.$$

The dynamics of V are obtained by using Itô's lemma for the Poisson process, as presented in Equation (5.9). There, we set $g(t, S(t)) := V(t, S)$ and $\bar{J}(t, S(t)) := (\mathrm{e}^J - 1)S$, which implies the following dynamics:

$$\begin{aligned} \mathrm{d}V = &\left(\frac{\partial V}{\partial t} + \bar{\mu}(t, S)\frac{\partial V}{\partial S} + \frac{1}{2}\bar{\sigma}^2(t, S)\frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t + \bar{\sigma}(t, S)\frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{Q}}(t) \\ &+ \left(V(t, S\mathrm{e}^J) - V(t, S)\right)\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t). \end{aligned}$$

After substitution, the dynamics of V/M read:

$$\begin{aligned} \mathrm{d}\frac{V}{M} = &\frac{1}{M}\left(\frac{\partial V}{\partial t} + \bar{\mu}(t, S)\frac{\partial V}{\partial S} + \frac{1}{2}\bar{\sigma}^2(t, S)\frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t + \frac{\bar{\sigma}(t, S)}{M}\frac{\partial V}{\partial S}\mathrm{d}W^{\mathbb{Q}} \\ &+ \frac{1}{M}\left(V(t, S\mathrm{e}^J) - V(t, S)\right)\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}} - r\frac{V(t, S)}{M}\mathrm{d}t. \end{aligned}$$

The jumps are independent of Poisson process $X_{\mathcal{P}}^{\mathbb{Q}}$ and Brownian motion $W^{\mathbb{Q}}$. Because V/M is a martingale, it follows that,

$$\begin{aligned} &\left(\frac{\partial V}{\partial t} + \bar{\mu}(t, S)\frac{\partial V}{\partial S} + \frac{1}{2}\bar{\sigma}^2(t, S)\frac{\partial^2 V}{\partial S^2} - rV\right)\mathrm{d}t \\ &\quad + \mathbb{E}\left[\left(V(t, S\mathrm{e}^J) - V(t, S)\right)\right]\mathbb{E}\left[\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}\right] = 0. \end{aligned} \tag{5.15}$$

Based on Equation (5.3) and substitution of the expressions (5.13) in (5.15), the following pricing equation results:

$$\frac{\partial V}{\partial t} + \left(r - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right]\right) S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - (r + \xi_p)V + \xi_p \mathbb{E}\left[V(t, S\mathrm{e}^J)\right] = 0, \tag{5.16}$$

which is a *partial integro-differential equation* (PIDE), as we deal with partial derivatives, but the expectation gives rise to an integral term.

PIDEs are typically more difficult to be solved than PDEs, due to the presence of the additional integral term. For only a few models analytic solutions exist. An analytic expression has been found for the solution of (5.16) for Merton's model and Kou's model, the solution is given in the form of an infinite series [Kou, 2002; Merton, 1976], see Section 5.2.1 to follow. However, in the literature a number of numerical techniques for solving PIDEs are available (see for e.g. [He *et al.*, 2006; Kennedy *et al.*, 2009]).

Example 5.1.2 For the jump diffusion process under measure $\mathbb{Q}$, we arrive at the following option valuation PIDE, in terms of the prices S,

$$\begin{cases} -\dfrac{\partial V}{\partial t} = \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 V}{\partial S^2} + (r - \xi_p \mathbb{E}[\mathrm{e}^J - 1]) S \dfrac{\partial V}{\partial S} - (r + \xi_p)V \\ \qquad + \xi_p \displaystyle\int_0^\infty V(t, S\mathrm{e}^y)\mathrm{d}F_J(y), \quad \forall (t, S) \in [0, T) \times \mathbb{R}_+, \\ V(T, S) = \max(\bar{\alpha}(S(T) - K), 0), \quad \forall S \in \mathbb{R}_+, \end{cases}$$

with $\bar{\alpha} = \pm 1$ (call or put, respectively), and where $\mathrm{d}F_J(y) = f_J(y)\mathrm{d}y$.

In log-coordinates $X(t) = \log(S(t))$, the corresponding PIDE for $V(t, X)$ is given by

$$\begin{cases} -\dfrac{\partial V}{\partial t} = \dfrac{1}{2}\sigma^2 \dfrac{\partial^2 V}{\partial X^2} + (r - \dfrac{1}{2}\sigma^2 - \xi_p \mathbb{E}[\mathrm{e}^J - 1]) \dfrac{\partial V}{\partial X} - (r + \xi_p)V \\ \qquad + \xi_p \displaystyle\int_{\mathbb{R}} V(t, X + y)\mathrm{d}F_J(y), \quad \forall (t, X) \in [0, T) \times \mathbb{R}, \\ V(T, X) = \max\left(\bar{\alpha}(\exp\left(X(T)\right) - K), 0\right), \quad \forall X \in \mathbb{R}. \end{cases}$$

Notice the differences in the integral terms of the two PIDEs above. ♦

5.1.3 Special cases for the jump distribution

Depending on the cumulative distribution function for the jump magnitude $F_J(x)$, a number of jump models can be defined.

Two popular choices are:

⇒ *Classical Merton' s model* [Merton, 1976]: The jump magnitude J is normally distributed, with mean μ_J and standard deviation σ_J. So, $\mathrm{d}F_J(x) = f_J(x)\mathrm{d}x$, where

$$f_J(x) = \frac{1}{\sigma_J\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_J)^2}{2\sigma_J^2}\right). \tag{5.17}$$

⇒ *Non-symmetric double exponential* (Kou's model [Kou, 2002; Kou and Wang, 2004])

$$f_J(x) = p_1\alpha_1 \mathrm{e}^{-\alpha_1 x}\mathbb{1}_{\{x\geq 0\}} + p_2\alpha_2\mathrm{e}^{\alpha_2 x}\mathbb{1}_{\{x<0\}}, \tag{5.18}$$

where p_1, p_2 are positive real numbers so that $p_1 + p_2 = 1$. To be able to integrate e^x over the real line it is required to have $\alpha_1 > 1$ and $\alpha_2 > 0$, and we obtain the expression

$$\mathbb{E}[\mathrm{e}^{J_k}] = p_1\frac{\alpha_1}{\alpha_1 - 1} + p_2\frac{\alpha_2}{\alpha_2+1}. \tag{5.19}$$

Example 5.1.3 (Paths of jump diffusion process) In Figure 5.2 examples of paths for $X(t)$ in (5.11) and $S(t) = \mathrm{e}^{X(t)}$, as in (5.10), are presented. Here, the classical Merton model is used, $J \sim \mathcal{N}(\mu_J, \sigma_J^2)$, where the jumps are symmetric, as described in (5.17).

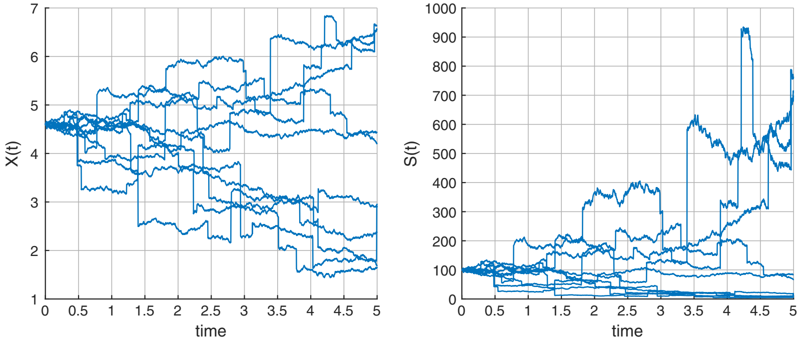


Figure 5.2: *Left side: Paths of process $X(t)$ (5.11); Right side: $S(t)$ in (5.10) with $S(t_0) = 100$, $r = 0.05$, $\sigma = 0.2$, $\sigma_J = 0.5$, $\mu_J = 0$, $\xi_p = 1$ and $T = 5$.*

5.2 Feynman-Kac theorem for jump diffusion process

The relation between the solution of a PDE and the computation of the discounted expected value of the payoff function, via the Feynman-Kac theorem (in Section 3.2), can be generalized to solving PIDEs that originate from the asset price processes $S(t)$ with jumps. As an example, we detail this for Merton's jump diffusion model.

With r constant, $S(t)$ is governed by the following SDE

$$\frac{\mathrm{d}S(t)}{S(t)} = \left(r - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right]\right)\mathrm{d}t + \sigma \mathrm{d}W^{\mathbb{Q}}(t) + \left(\mathrm{e}^J - 1\right)\mathrm{d}X_{\mathcal{P}}^{\mathbb{Q}}(t),$$

and option value $V(t,S)$ satisfies the following PIDE,

$$\begin{aligned}\frac{\partial V}{\partial t} &+ \left(r - \xi_p \mathbb{E}\left[\mathrm{e}^J - 1\right]\right) S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - (r+\xi_p)V \\ &+ \xi_p \mathbb{E}\left[V(t, S\mathrm{e}^J)\right] = 0,\end{aligned}$$

with $V(T,S) = H(T,S)$, and the term $(\mathrm{e}^J - 1)$ representing the size of a proportional jump, the risk-neutral valuation formula determines the option value, i.e.

$$V(t,S) = M(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{M(T)}H(T,S)|\mathcal{F}(t)\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-r(T-t)}H(T,S)|\mathcal{F}(t)\right].$$

The discounted expected payoff formula resembles the well-known martingale property if we consider the quantity $\frac{V(t,S)}{M(t)}$, which should not contain a drift term, i.e. $\mathbb{E}^{\mathbb{Q}}\left[\mathrm{d}(V(t,S)/M(t))|\mathcal{F}(t)\right] = 0$. Equation (5.15) details this condition in the case of jump diffusion, confirming that then PIDE (5.16) should be satisfied. So, also in the case of this type of jump processes, we can solve the pricing PIDE in the form of the calculation of a discounted expected payoff, by means of the well-known Feynman-Kac theorem.

We will derive analytic option prices as well as the characteristic function for the jump diffusion process. Before we derive these, we present an application of the tower property of expectations for discrete random variables, as we will need this in the derivations to follow.

Example 5.2.1 (Tower property, discrete random variables) As an example of how convenient the conditional expectation can be used, we consider the following problem. Suppose $X_1, X_2, \ldots$ are independent random variables with the same mean μ and N_J is a nonnegative, integer-valued random variable which is independent of all X_i's. Then, the following equality holds:

$$\mathbb{E}\left[\sum_{i=1}^{N_J} X_i\right] = \mu \mathbb{E}\left[N_J\right]. \tag{5.20}$$

This equality is also known as *Wald's equation.*

Using the tower property (1.20) for discrete random variables z_1 and z_2, we find,

$$\mathbb{E}\left[\mathbb{E}\left[z_1 | z_2\right]\right] = \sum_z \mathbb{E}[z_1 | z_2 = z]\, \mathbb{P}[z_2 = z],$$

so that,

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{N_J} X_k\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{N_J} X_k \Big| N_J\right]\right] \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[\sum_{k=1}^{n} X_k \Big| N_J = n\right] \mathbb{P}[N_J = n] \\
&= \sum_{n=1}^{\infty} \mathbb{P}[N_J = n] \sum_{k=1}^{n} \mathbb{E}\left[X_k\right].
\end{aligned} \tag{5.21}$$

Since the expectation for each X_k equals μ, we have:

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{N_J} X_k\right] &= \sum_{n=1}^{\infty} \mathbb{P}[N_J = n] \sum_{k=1}^{n} \mu \\
&= \mu \sum_{n=1}^{\infty} n \mathbb{P}[N_J = n] \stackrel{\text{def}}{=} \mu \mathbb{E}[N_J],
\end{aligned} \tag{5.22}$$

which confirms (5.20). ♦

5.2.1 Analytic option prices

In this section we derive the solution to call option prices for an asset under the risk-neutral *standard jump diffusion model*, as defined in Equation (5.10). Integrating the log-transformed asset price process in Equation (5.11), we find:

$$\begin{aligned}
X(T) &= X(t_0) + \int_{t_0}^{T} \left(r - \xi_p \mathbb{E}[\mathrm{e}^J - 1] - \frac{1}{2}\sigma^2\right) \mathrm{d}t \\
&\quad + \int_{t_0}^{T} \sigma \mathrm{d}W(t) + \int_{t_0}^{T} J(t) \mathrm{d}X_{\mathcal{P}}(t) \\
&= X(t_0) + \left(r - \xi_p \mathbb{E}[\mathrm{e}^J - 1] - \frac{1}{2}\sigma^2\right)(T - t_0) \\
&\quad + \sigma\left(W(T) - W(t_0)\right) + \sum_{k=1}^{X_{\mathcal{P}}(T)} J_k,
\end{aligned}$$

where $S(t) = \mathrm{e}^{X(t)}$.