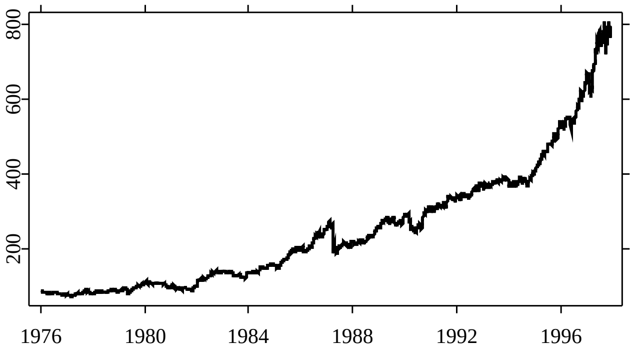
Introduction to Statistics of Financial Markets



Jürgen Franke
Wolfgang Härdle
Christian Hafner

Please note: this is only a sample of the full book. The complete book can be downloaded on the e-book page of XploRe. Just click the download logo:

http://www.xplore-stat.de/ebooks/ebooks.html

download logo

For further information please contact MD*Tech at mdtech@mdtech.de

Jürgen Franke
Universität Kaiserslautern
Fachbereich Mathematik
Postfach 3049
67653 Kaiserslautern
franke@mathematik.uni-kl.de

Wolfgang Härdle
Humboldt-Universität zu Berlin
CASE - Center for Applied Statistics and Economics
Institut für Statistik und Ökonometrie
10178 Berlin
haerdle@wiwi.hu-berlin.de

Christian Hafner
Erasmus Universiteit Rotterdam
Faculteit der Economische Wetenschappen (FEW)
Postbus 1738
3000 DR Rotterdam
chafner@few.eur.nl

Figure 0.1: Notes of a student for the exam of a course based on this book.

Figure 0.2: Notes of a student for the exam of a course based on this book.

Figure 0.3: Notes of a student for the exam of a course based on this book.

Figure 0.4: Notes of a student for the exam of a course based on this book.

Figure 0.5: Notes of a student for the exam of a course based on this book.

Preface

Until about the 1970s, financial mathematics has been rather modest compared with other mathematical disciplines. This changed rapidly after the path-breaking works of F. Black, M. Scholes, and R. Merton on derivative pricing, for which they received the Nobel prize of economics in 1997. Since 1973, the publication year of the famous Black and Scholes article, the importance of derivative instruments in financial markets has not ceased to grow. Higher risks associated with, for example, flexible instead of fixed exchange rates after the fall of the Bretton Woods system required a risk management and the use of hedging instruments for internationally active companies. More recently, globalization and the increasingly complex dependence of financial markets are reasons for using sophisticated mathematical and statistical methods and models to evaluate risks.

The necessity to improve and develop the mathematical foundation of existing risk management was emphasized in the turbulent 1990s with, for example, the Asian crisis, the hedging disasters of Metallgesellschaft and Orange County, and the fall of the Long-Term Capital Management hedge fund (controlled by Merton and Scholes!). This saw the legislator obliged to take action. In continental Europe, this development is mainly influenced by the Basel Committee on Banking Supervision, whose recommendations form the basis in the European Union for legislation, with which financial institutions are obliged to do a global, thorough risk management. As a result, there is an increasing demand for experts in financial engineering, who control risks internally, search for profitable investment opportunities and guarantee the obligations of legislation. In the future, such risk management is likely to become obligatory for other, deregulated markets such as telecommunication and energy markets. Being aware of the increasing price, volume, and credit risks in these markets, large companies usually have already created new departments dealing with asset and liability management as well as risk management.

The present text is supposed to deliver the necessary mathematical and statistical basis for a position in financial engineering. Our goal is to give a comprehensive introduction into important ideas of financial mathematics and statistics. We do not aim at covering all practically relevant details, and we also do not discuss the technical subtleties of stochastic analysis. For both purposes there is already a vast variety of textbooks. Instead, we want to give students of mathematics, statistics, and economics a primer for the modelling and statistical analysis of financial data. Also, the book is meant for practitioners, who want to deepen their acquired practical knowledge. Apart from an introduction to the theory of pricing derivatives, we emphasize the statistical aspects of mathematical methods, i.e., the selection of appropriate models as well as fitting and validation using data.

The present book consists of three parts. The first two are organized such that they can be read independently. Each one can be used for a course of roughly 30 hours. We deliberately accept an occasional redundancy if a topic is covered in both parts but from a different perspective. The third part presents selected applications to current practical problems. Both *option pricing* as *statistical modelling of financial time series* have often been topic of seminars and lectures in the international study program *financial mathematics* of Universität Kaiserslautern (www.mathematik.uni-kl.de) as well as in the economics and statistics program of Humboldt-Universität zu Berlin (ise.wiwi.hu-berlin.de). Moreover, they formed the basis of lectures for banking practitioners which were given by the authors in various European countries.

The first part covers the classical theory of pricing derivatives. Next to the Black and Scholes option pricing formula for conventional European and American options and their numerical solution via the approximation using binomial processes, we also discuss the evaluation of some exotic options. Stochastic models for interest rates and the pricing of interest rate derivatives conclude the first part. The necessary tools of stochastic analysis, in particular the Wiener process, stochastic differential equations and Itô's Lemma will be motivated heuristically and not derived in a rigorous way. In order to render the text accessible to non-mathematicians, we do not explicitly cover advanced methods of financial mathematics such as martingale theory and the resulting elegant characterization of absence of arbitrage in complete markets.

The second part presents the already classical analysis of financial time series, which originated in the work of T. Bollerslev, R. Engle, and C. Granger. Starting with conventional linear processes, we motivate why financial time series rarely can be described using such linear models. Alternatively, we discuss the related model class of stochastic volatility models. Apart from standard ARCH and GARCH models, we discuss extensions that allow for an asymmetric impact of lagged returns on volatility. We also review multivariate GARCH models that can be applied, for example, to estimate and test the capital asset pricing model (CAPM) or to portfolio selection problems. As a support for explorative data analysis and the search and validation of parsimonious parametric models, we emphasize the use of nonparametric models for financial time series and their fit to data using kernel estimators or other smoothing methods.

In the third part of the book, we discuss applications and practical issues such as option pricing, risk management, and credit scoring. We apply flexible GARCH type models to evaluate options and to overcome the Black and Scholes restriction of constant volatility. We give an overview of Value at Risk (VaR) and backtesting, and show that copulas can improve the estimation of VaR. A correct understanding of the statistical behavior of extremes such as September 11, 2001, is essential for risk management, and we give an overview of extreme value theory with financial applications. As a particularly popular nonparametric modelling tool in financial institutions, we discuss neural networks from a statistical viewpoint with applications to the prediction of financial time series. Next, we show how a principal components analysis can be used to explain the dynamics of implied volatilities. Finally, we present nonparametric extensions of conventional discrete choice models and apply them to the credit scoring problem.

We decided to collect some technical results concerning stochastic integration in the appendix. Here we also present Girsanov's theorem and the martingale representation theorem, with which dynamic portfolio strategies as well as an alternative proof of the Black and Scholes formula are developed. This appendix is based on work by Klaus Schindler, Saarbrücken.

In designing the book as e-book, we are going new ways of scientific publishing together with Springer Verlag and MD*Tech. The book is provided with an individual license key, which enables the reader to download the html and pdf versions of the text as well as all slides for a 60 to 90 hours lecture from the e-book server at www.quantlet.com. All examples, tables and graphs can be reproduced and changed interactively using the XploRe quantlet technology.

The present book would not exist without the cooperating contributions of P. Čížek, M. Fengler, Z. Hlávka, E. Kreutzberger, S. Klinke, D. Mercurio and D. Peithmann. The first part of the book arose from an extended vocational training which was developed together with G. Maercker, K. Schindler and N. Siedow. In particular, we want to thank Torsten Kleinow, who accompanied the writing of the text in all phases, developed the e-book platform and improved the presentation by various valuable contributions. Important impulses for an improved presentation were given by Klaus Schindler of the

University of Saarbrücken, which we gratefully acknowledge. The chapter on copulas is based on a contribution by Jörn Rank, Andersen Consulting, and Thomas Siegl, BHF Bank, which we adopted with their kind approval. The quantlets for multivariate GARCH models were contributed by Matthias Fengler and Helmut Herwartz. All graphs were created by Ying Chen, who also led the text management. We would like to express our thanks to these colleagues. We also benefitted from many constructive comments by our students of the universities in Kaiserslautern, Berlin, and Rotterdam. As an example of their enthusiasm we depict the preparation sheet of a student for the exam at the front pages of the book. Graphs and formulae are combined to create a spirit of the "art of quantitative finance".

Finally, for the technical realization of the text we want to thank Beate Siegler and Anja Ossetrova.

Kaiserslautern, Berlin and Rotterdam, April 2004

Frequently Used Notation

$x \stackrel{\mathrm{def}}{=} \ldots$ x is defined as ...
$\mathbb{R}$ real numbers
$\overline{\mathbb{R}} \stackrel{\mathrm{def}}{=} \mathbb{R} \cup \{\infty, \infty\}$
$A^\top$ transpose of matrix A
$X \sim D$ the random variable X has distribution D
$\mathsf{E}[X]$ expected value of random variable X
$Var(X)$ variance of random variable X
$\mathrm{Cov}(X, Y)$ covariance of two random variables X and Y
$\mathrm{N}(\mu, \Sigma)$ normal distribution with expectation μ and covariance matrix Σ, a similar notation is used if Σ is the correlation matrix
$\mathrm{P}[A]$ or $\mathrm{P}(A)$ probability of a set A
$\mathbf{1}$ indicator function
$(F \circ G)(x) \stackrel{\mathrm{def}}{=} F\{G(x)\}$ for functions F and G
$x \approx y$ x is approximately equal to y
$\alpha_n = \mathcal{O}(\beta_n)$ iff $\frac{\alpha_n}{\beta_n} \longrightarrow$ constant, as $n \longrightarrow \infty$
$\alpha_n = \mathcal{o}(\beta_n)$ iff $\frac{\alpha_n}{\beta_n} \longrightarrow 0$, as $n \longrightarrow \infty$
$\mathcal{F}_t$ is the information set generated by all information available at time t
Let A_n and B_n be sequences of random variables.
$A_n = \mathcal{O}_p(B_n)$ iff $\forall \varepsilon > 0\ \exists M,\ \exists N$ such that $\mathrm{P}[|A_n/B_n| > M] < \varepsilon,\ \forall n > N$.
$A_n = \mathcal{o}_p(B_n)$ iff $\forall \varepsilon > 0\ :\ \lim_{n\to\infty} \mathrm{P}[|A_n/B_n| > \varepsilon] = 0$.

Contents

Part I

Option Pricing

1 Derivatives

Classical financial mathematics deals first of all with basic financial instruments like stocks, foreign currencies and bonds. A *derivative* (*derivative security* or *contingent claim*) is a financial instrument whose value depends on the value of others, more basic *underlying* variables. In this chapter we consider forward contracts, futures contracts and options as well as some combinations.

Simple derivatives have been known on European stock exchanges since the turn of the 19th century. While they lost popularity between World War I and II, they revived in the seventies accompanied by the work of Black, Scholes and Merton, who developed a theoretical foundation to price such instruments. Their entrepreneurial approach, which is not only applied to price derivatives but everywhere in finance where the risk of complex financial instruments is measured and controlled, was awarded the Nobel price in economics in 1997. At the same time, it triggered the development of modern financial mathematics whose basics we describe in the first part of this book. Since we concentrate only on the mathematical modelling ideas, we introduce the required financial terminology as we pass by. We leave out numerous details which are of practical importance but which are of no interest for the mathematical modelling, and refer to, for example, Hull (2000), Welcker, Kloy and Schindler (1992).

Particularly simple derivative securities are *forward* and *future contracts*. Both contracts are agreements involving two parties and calling for future delivery of an asset at an agreed–upon price. Stocks, currencies and bonds, as well as agricultural products (grain, meat) and raw materials (oil, copper, electric energy) are underlying in the contract.

Definition 1.1 (Forward contract)
A forward contract *is an agreement between two parties in which one of the parties assumes a long position (the other party assumes a short position) and obliges to purchase (sell) the underlying asset at a specified future date $T > t$, (*expiration date *or* maturity*) for a specified price K (*delivery price*).*

At time t, the value $V_{K,T}(S_t, \tau)$ *of such a contract depends on the current value of the underlying* S_t*, the time to maturity* $\tau = T - t$ *and of the parameters* K*,* T *specified in the contract.*

Futures contracts closely resemble forward contracts. While the latter do not entail any more payments until maturity once the agreement is signed, futures contracts are traded on an exchange and mark to the market on a daily basis. Under certain assumptions forward and futures prices are identical.

Example 1.1
An investor enter into a long forward contract on September 1, 2003, which obliges him to buy 1 000 000 EUR *at a specified exchange rate of 1.2* USD/EUR *in 90 days. The investor gains if the exchange rate is up to 1.3* USD/EUR *on November 30, 2003. Since he can sell the 1 000 000* EUR *for* USD *1 300 000. In this case t = September 1, 2003* τ *= 90 days, T = November 30, and* $K =$ USD *1 200 000.*

Definition 1.2 (Spot Price, Forward Price, Future Price)
The current price of the underlying (stock, currency, raw material) S_t *is often referred to as the* spot price. *The delivery price giving a forward contract a value of zero is called* forward price *and denoted* F_t*. That is,* F_t *solves* $V_{F_t,T}(S_t, \tau) = 0$*. The* future price *is defined accordingly.*

Later we will compute the value of a forward contract, which determines the forward price. Since under certain assumptions forward and future contracts have the same value, their prices are equal. When such a contract is initiated in time $t = 0$, often the delivery price is set to $K = F_0$. The contract has a value of zero for both the seller and the buyer, i.e. no payments occur. In the course of time, as additional transactions take place on the exchange, the delivery price K and the forward price F_t can be different.

Contrary to forward and futures contracts where both parties are obligated to carry out the transaction, an option gives one party the right to buy or sell the security. Obviously, it's important to distinguish whether the buyer or seller of the option has the right to exercise. There are two types of options: call and put options. Furthermore, European options are delimited from American options. While European options are like forward contracts, American options can be exercised at any date before maturity. These terms are derived from historical, not geographical roots.

Definition 1.3 (Call Option, Put Option)
A European call option *is an agreement which gives the holder the right to buy*

*the underlying asset at a specified date $T > t$, (*expiration date *or* maturity*), for a specified price K, (*strike price *or* exercise price*). If the holder does not exercise, the option expires worthless.*

European put option *is an agreement which gives the holder the right to sell the underlying asset at a specified date T for a specified price K.*

The holder of an American call *or* put option *has the right to exercise the option at any time between t and T.*

The option types defined above are also called *plain vanilla options*. In practice, many more complex derivatives exist and numerous new financial instruments are still emerging. *Over-the-counter (OTC) derivatives* are tailor made instruments designed by banking institutions to satisfy a particular consumer need. A compound option, for example, is such an OTC–derivative. It gives the holder the right to buy or sell at time T an underlying option which matures in $T' > T$. The mathematical treatment of these *exotic options* is particularly difficult, since the current value of this instrument does not only depend on the value of the underlying S_t but also on the entire path of the underlying, $S_{t'}, 0 \leq t' \leq t$.

Asian, lookback and knock–out options are path-dependent derivatives. While the delivery price K of an asian option depends on the average value of the security of a certain period of time, it depends in the case of a lookback option on the minimum or maximum value of the security for a certain period of time. Knock–out options expire worthless if the price level ever hits a specified level.

To get used to forward and futures contracts, plain vanilla options and simple combinations of them, it is convenient to have a look at the *payoff* of an instrument, i.e. the value of the derivative at maturity T. The payoff of a long position in a forward contract is just $S_T - K$, with S_T the security's spot price at expiration date T. The holder of the contract pays K for the security and can sell it for S_T. Thus, he makes a profit if the value of the security S_T at expiration is greater than the delivery price K. Being short in a forward contract implies a payoff $K - S_T$. Both payoff functions are depicted in Figure 1.1.

The call option payoff function is denoted:

$$\max\{S_T - K, 0\} = (S_T - K)^+.$$

Thus, the option holder only exercises if the delivery price K is less than the value of the security S_T at expiration date T. In this case, he receives the same cash amount as in the case of a forward or future contract. If $K < S_T$,

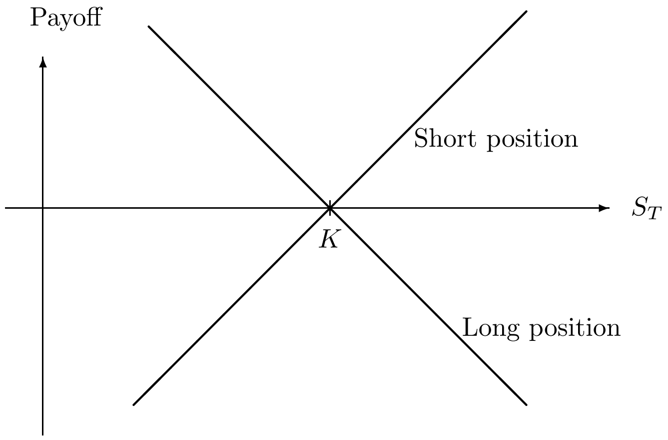


Figure 1.1: Value of forward contract at maturity

he will clearly choose not to exercise and the option expires worthless. The put option payoff function is:

$$\max\{K - S_T, 0\} = (K - S_T)^+.$$

In contrast to forward and future contracts, options have to be bought for a positive amount $C(S_0, T)$, called the *option price* or *option prime*. Often, the options profit function is defined as $(S_T - K)^+ - C(S_0, T)$. However, this definition adds cash flows of different points in time. The correct profit is obtained by compounding the cash outflow in time $t = 0$ up to time $t = T$, since the investor could have invested the option option at the risk–free interest rate r. Assuming continuous compounding at a constant interest rate r, the profit function of a call option is denoted: $(S_T - K)^+ - C(S_0, T)e^{rT}$.

Example 1.2
Consider a long call option with delivery price K and option price C_0 in time $t = 0$. The payoff and profit function are given in Figure 1.2 and 1.3, respectively.

Example 1.3
Combining a long call and a long put with the same delivery price K is called

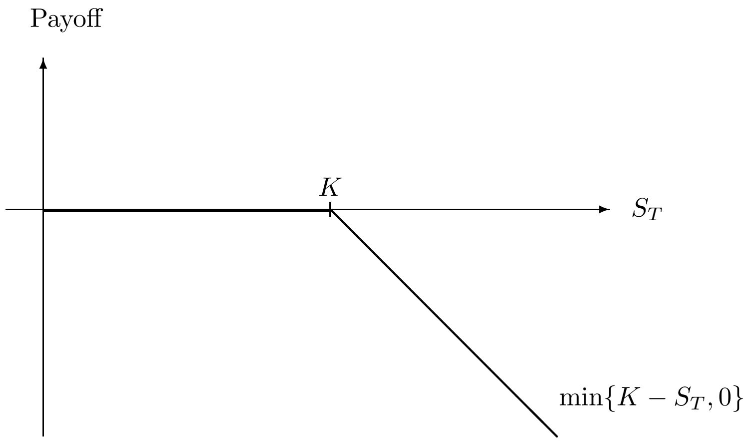


Figure 1.2: Payoff of a short position in a call option

a straddle. *Figure 1.4 shows the straddle profit function.* C_0 *and* P_0 *denote the call and put option option respectively.*

Another fundamental financial instrument which is used in option pricing is a *bond*. Apart from interest yields, the bond holder possibly receives coupon payments at fixed points in time. In particular, we will consider zero–coupon bonds, i.e. bonds which promise a single payment at a fixed future date.

Definition 1.4 (Zero coupon Bond, Discount Bond)
A zero coupon bond *or* discount bond *is a bond without coupon payments which pays an interest* r. *The investor pays in time 0 an amount* B_0 *and receives at maturity* T *the amount* B_T *which is the sum of* B_0 *and the interest earned on* B_0. *The bonds' value at maturity is termed* face value.

Buying a zero–coupon bond corresponds to lending money at a fixed interest rate for a fixed period of time. Conversely, selling a zero-coupon bond is equivalent to borrowing money at rate r. Since bonds are traded on an exchange, they can be sold prior to maturity at price B_t, i.e. B_0 plus accrued interest up to time t.

In practice, interest rates are compounded at discrete points in time, for

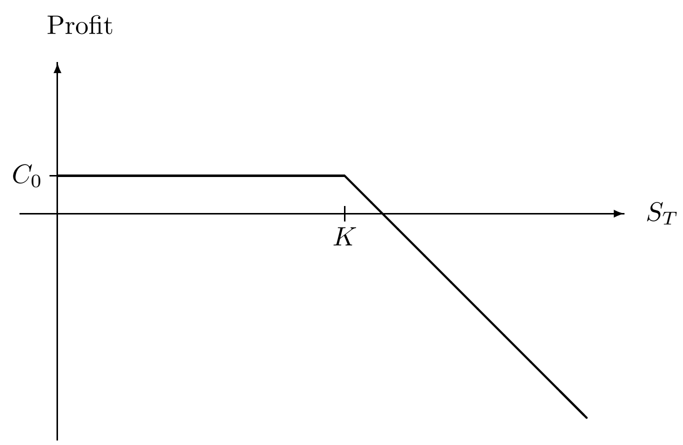


Figure 1.3: Profit of a short position in a call option

example annually, semiannually or monthly. If the interest rate r is compounded annually, the initial investment B_0 has n years later a value of $B_n^{(1)} = B_0(1+r)^n$. If it is compounded k times per annum (p.a.), the investment pays an interest rate of $\frac{r}{k}$ each $\frac{1}{k}$ years, and has a terminal value of $B_n^{(k)} = B_0(1+\frac{r}{k})^{nk}$ after n years. However, when options and other complex derivatives are priced, continuous compounding is used, which denoted for $k \to \infty$. In this case, the initial investment B_0 grows in n years to $B_n = B_0 \cdot e^{nr}$, and r is called *short rate*. The difference between discrete and continuous compounding is small when k is large. While an investment of $B_0 = 1000$ EUR at a yearly rate $r = 10\%$ grows to 1100 EUR within a year when annually compounded, it grows to 1105.17 EUR when continuously compounded.

In light of this, the continuous compounded rate r can be modified to account for these deviations. Assuming annual compounding at rate r_1, for both continuous and annual compounding, a continuous compounded rate $r = \log(1+r_1)$ has to be applied, in order to obtain the same terminal value $B_n = B_n^{(1)}$.

If not stated otherwise, continuous compounding will be assumed from here on. For comparing cash flows occurring at different points in time, they

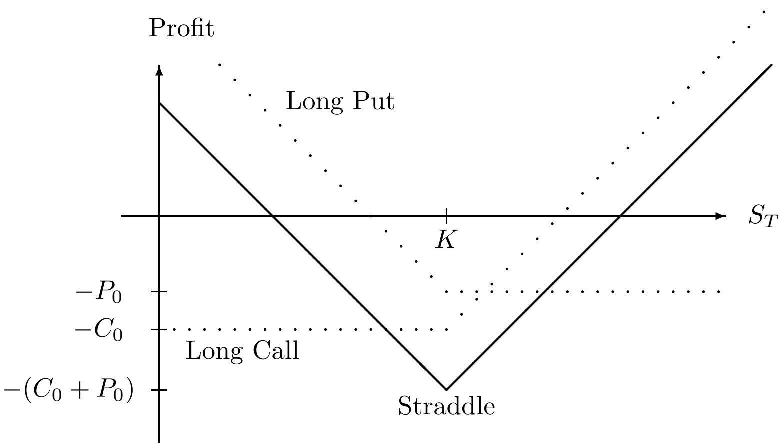


Figure 1.4: Profit of a straddle

have to be compounded or discounted to the same point in time. That is, interest payments are added or subtracted. With continuous compounding, an investment of B in time t in $\Delta t > 0$ is

compounded to time $t + \Delta t$: $B\, e^{r\Delta t}$

discounted to time $t - \Delta t$: $B\, e^{-r\Delta t}$.

Before finishing the chapter, some more financial terms will be introduced. A *portfolio* is a combination of one or more financial instruments - its value is considered as an individual financial instrument. One element of a portfolio is also called a *position*. An investor assumes a *long position* when he buys an instrument, and a *short position* when he sells it. A *long call* results from buying a call option, a *long put* from buying a put option, and *a short forward* from selling a forward contract.

An investor closes out a position of his portfolio by making the future portfolio performance independent of the instrument. If the latter is traded on an exchange, he can sell (e.g. a stock or a bond) or buy (e.g. borrowed money) it. Should the instrument not be traded, however, the investor can close out the position by adding to the portfolio the inverse instrument. Thus, both sum up to zero, and do not influence the portfolio performance any more.

Example 1.4
Consider an investor who bought on February 1 a 1 000 000 USD *forward contract with a delivery price of 1 200 000* EUR *and which matures in one year. On June 1, he wishes to close out the position. He can* sell *another forward contract of the same size with the same delivery price and the maturity date, namely January 31. The long and the short positions sum up to zero at any time.*

Short selling is a trading strategy that involves selling financial instruments, for example stocks, which he does not own. At a later point in time, he buys back these objects. In practice, this requires the intervention of a broker who mediates another client owing the objects and willing to lend them to the investor. The short selling investor commits to pay to the client any foregone income, as dividends for example, that would be received in the meantime.

Example 1.5
An investor selling short 1000 stocks, lends them from the owner and sells them immediately for 1000 S_0 *in the market (*S_t *denotes the stock price at time* t*). Later, at time* $t > 0$*, he closes out the position, by buying back the stocks for 1000* S_t *and returning them to the owner. The strategy is profitable if* S_t *is clearly below* S_0*. If in time* t_0*,* $0 < t_0 < t$*, a dividend* D *per share is paid, the investor pays 1000* D *to the owner. Short selling is in practice subject to numerous restrictions. In the following, it is only the possibility of short selling that will be of interest.*

1.1 Recommended Literature

Basic textbooks on derivatives are, among others, Hull (2000), Jarrow (1992) and Cox and Rubinstein (1985). Neftci (1996) and Duffie (1996) are more advanced regarding the mathematical level. A rather practical but still theoretically well–founded introduction is given by Briys, Bellalah, Mai and de Varenne (1998).

2 Introduction to Option Management

2.1 Arbitrage Relations

In this section we consider the fundamental notion of no–arbitrage. An *arbitrage opportunity* arises if it is possible to make a riskless profit. In an ideal financial market, in which all investors dispose of the same pieces of information and in which all investors can react instantaneously, there should not be any arbitrage opportunity. Since otherwise each investor would try to realize the riskless profit instantaneously. The resulting transactions would change the prices of the involved financial instruments such that the arbitrage opportunity disappears.

Additionally to no–arbitrage we presume in the remaining chapter that the financial market fulfills further simplifying assumptions which are in this context of minor importance and solely serve to ease the argumentation. If these assumptions hold we speak of a perfect financial market.

ASSUMPTION *(perfect financial market)*
There are no arbitrage opportunities, no transaction costs, no taxes, and no restrictions on short selling. Lending rates equal borrowing rates and all securities are perfectly divisible.

The assumption of a perfect financial market is sufficient to determine the value of future and forward contracts as well as some important relations between the prices of some types of options. Above all no mathematical model for the price of the financial instrument is needed. However, in order to determine the value of options more than only economic assumptions are necessary. A detailed mathematical modelling becomes inevitable. Each mathematical approach though has to be in line with certain fundamental arbitrage relations being developed in this chapter. If the model implies values of future and forward contracts or option prices which do not fulfill these relations the model's assumptions must be wrong.

An important conclusion drawn from the assumption of a perfect financial market and thus from no–arbitrage will be used frequently in the proofs to come. It is the fact that two portfolios which have at a certain time T the same value must have the same value at a prior time $t < T$ as well. Due to its importance we will further illustrate this reasoning. We proceed from two portfolios A and B consisting of arbitrary financial instruments. Their value in time t will be denoted by $W_A(t)$ and $W_B(t)$ respectively. For any fixed point of time T, we assume that $W_A(T) = W_B(T)$ independently of the prior time T values of each financial instrument contained in A and B. For any prior point of time $t < T$ we assume without loss of generality that $W_A(t) \leq W_B(t)$. In time t an investor can construct without own financial resources a portfolio which is a combination of A and B by buying one unit of every instrument of A, selling one unit of every instrument of B (short selling) and by investing the difference $\Delta(t) = W_B(t) - W_A(t) \geq 0$ at a fixed rate r. The combined portfolio has at time t a value of

$$W_A(t) - W_B(t) + \Delta(t) = 0,$$

i.e. the investor has no initial costs. At time T the part of the combined portfolio which is invested at rate r has the compounded value $\Delta(T) = \Delta(t)e^{r(T-t)}$, and hence the combined portfolio has a value of

$$W_A(T) - W_B(T) + \Delta(T) = \Delta(t)e^{r(T-t)} > 0,$$

if $\Delta(t) > 0$. The investor made a riskless gain by investing in the combined portfolio which contradicts the no–arbitrage assumption. Therefore, it must hold $\Delta(t) = 0$, i.e. $W_A(t) = W_B(t)$.

The previous reasoning can be used to determine the unknown value of a financial derivative. For this, a portfolio A is constructed which contains instruments with known price along with one unit of the derivative under investigation. Portfolio A will be compared to another portfolio B, called the *duplicating portfolio*, which contains exclusively instruments with known prices. Since the duplicating portfolio B is constructed such that for certain it has the same value at a fixed point of time T as portfolio A the no–arbitrage assumption implies that both portfolios must have the same value at any prior point of time. The value of the financial derivative can thus be computed at any time $t \leq T$. We illustrate this procedure in the following example of a forward contract.

Theorem 2.1
We consider a long forward contract to buy an object which has a price of S_t at time t. Let K be the delivery price, and let T be the maturity date. $V(s, \tau)$

denotes the value of the long forward contract at time t as a function of the current price $S_t = s$ and the time to maturity $\tau = T - t$. We assume constant interest rates r during the time to maturity.

1. *If the underlying object does not pay any dividends and does not involve any costs during the time to maturity τ, then it holds*

$$V(S_t, \tau) = V_{K,T}(S_t, \tau) = S_t - Ke^{-r\tau} \tag{2.1}$$

   The forward price is equal to $F_t = S_t e^{r\tau}$.

2. *If during the time to maturity the underlying pays at discrete time points dividends or involves any costs whose current time t discounted total value is equal to D_t, then it holds*

$$V(S_t, \tau) = V_{K,T}(S_t, \tau) = S_t - D_t - Ke^{-r\tau} \tag{2.2}$$

   The forward price is equal to $F_t = (S_t - D_t)e^{r\tau}$.

3. *If the underlying involves continuous costs at rate b, then it holds*

$$V(S_t, \tau) = V_{K,T}(S_t, \tau) = S_t e^{(b-r)\tau} - Ke^{-r\tau} \tag{2.3}$$

   The forward price is equal to $F_t = S_t e^{b\tau}$.

Proof:
For simplicity we assume the underlying object to be a stock paying either discrete dividend yields whose value discounted to time t is D_t or paying a continuous dividend yield at rate b. In the latter case the stock involves continuous costs equal to $b = r - d$. The investor having a long position in the stock gains dividends (as negative costs) at rate d but simultaneously loses interests at rate r since he invested his capital in the stock instead of in a bond with a fixed interest rate. In place of stocks, bonds, currencies or other simple instruments can be considered as well.

1. We consider at time t the following two portfolios A and B:

Portfolio A: One long forward contract on a stock with delivery price K, maturing in time T.
One long zero bond with face value K, maturing in time T.

Portfolio B: A long position in one unit of the stock.

At maturity T portfolio A contains a zero bond of value K. Selling this zero bond for K the obligation to buy the stock for K can be fulfilled. Following these transactions portfolio A consists as well as portfolio B of one unit of the stock. Thus both portfolios have at time T the same value and must therefore, due to the no–arbitrage assumption, have the same value at any time t prior to T :

$$V(S_t, \tau) + Ke^{-r\tau} = S_t \ , \tag{2.4}$$

since the value of the zero bond at time t is given by discounting K at rate r, $Ke^{-r\tau}$. The forward price is by definition the solution of

$$0 = V_{F_t,T}(S_t, \tau) = S_t - F_t e^{-r\tau}.$$

2. We consider at time t the two portfolios A and B as given above and add one position to portfolio B :

Portfolio B**:** A long position in one unit of the stock and one short position of size D_t in a zero bond with interest rate r (lending an amount of money of D_t).

At maturity T the dividend yields of the stock in portfolio B, which compounded to time T amount to $D_t e^{r\tau}$, are used to pay back the bond. Thus, both portfolios A and B consist again of one unit of the stock, and therefore they must have the same value at any time $t < T$:

$$V(S_t, \tau) + Ke^{-r\tau} = S_t - D_t \ . \tag{2.5}$$

The forward price results as in part 1 from the definition.

3. If the stock pays dividends continuously at a rate d, then the reasoning is similar as in part 2. Once again, we consider at time t two portfolios A and B. And again, A is left unchanged, B is now composed of the following position:

Portfolio B**:** A long position in $e^{-d\tau}$ stocks.

Reinvesting the dividends yields continuously in the stock portfolio B consists again of exactly one stock at time T. Heuristically, this can be illustrated as follows: In the time interval $[t, t+\delta]$ the stock pays approximately, for a small δ, a dividend of $d \cdot \delta \cdot S_t$. Thus, the current total amount of stocks in the portfolio, $e^{-d\tau} = e^{-d(T-t)}$, pays a total dividend yield of $d \cdot \delta \cdot S_t \cdot e^{-d(T-t)}$, which is reinvested in the stock. Assuming that the stock price does not change significantly in the interval $[t, t+\delta]$, i.e. $S_{t+\delta} \approx S_t$, portfolio B contains in time $t+\delta$

$$(1 + d \cdot \delta) \cdot e^{-d(T-t)} \approx e^{d\delta} \cdot e^{-d(T-t)} = e^{-d(T-t-\delta)}$$

stocks. The above reasoning can be done exactly by taking the limit $\delta \to 0$, and it can be shown that portfolio B contains at any time s between t and T exactly $e^{-d(T-s)}$ stocks. That is, for $s = T$ portfolio B is composed of exactly one stock. The same reasoning as in part 1 leads to the conclusion that portfolio A and B must have the same value at any time t. Thus, we have

$$V(S_t, \tau) + Ke^{-r\tau} = e^{-d\tau} S_t \; . \tag{2.6}$$

where we have to set $b = r - d$. The forward price results as in part 1 from the definition. □

Example 2.1 *We consider a long forward contract on a 5 year bond which is currently traded at a price of 900* EUR. *The delivery price is 910* EUR, *the time to maturity of the forward contract is one year. The coupon payments of the bond of 60* EUR *occur after 6 and 12 months (the latter shortly before maturity of the forward contract). The continuously compounded annual interest rates for 6 and 12 months are 9% and 10% respectively. In this example we have*

$$S_t = 900\; ,\; K = 910\; ,\; r = 0.10\; ,\; \tau = 1\; ,\; D_t = 60e^{-0.09\cdot\frac{1}{2}} + 60e^{-0.10} = 111.65 \tag{2.7}$$

Thus, the value of the forward contract is given by

$$V(S_t, \tau) = 900 - 111.65 - 910e^{-0.10} = -35.05. \tag{2.8}$$

The value of the respective short position in the forward contract is +35.05. The price F_t of the forward contract is equal to $F_t = (S_t - D_t)e^{r\tau} = 871.26$.

Example 2.2 *Consider a long forward contract to buy 1000 Dollar. If the investor buys the 1000 Dollar and invests this amount in a American bond, the American interest rate can be interpreted as a dividend yield d which is continuously paid. Let r be the home interest rate. The investment involves costs $b = r - d$, which are the difference between the American and the home interest rate. Denoting the dollar exchange rate by S_t the price of the forward contract is then given by*

$$F_t = S_t e^{b\tau} = S_t e^{(r-d)\tau}. \tag{2.9}$$

While for $r > d$ a report $S_t < F_t$ results, for $r < d$ a backwardation $S_t > F_t$ results. If $r > d$ and the delivery price is chosen to equal the current exchange rate, i.e. $K = S_t$, then the value of the forward contract is

$$V_{S_t,T}(S_t, \tau) = S_t(e^{-d\tau} - e^{-r\tau}) > 0.$$

Buying the forward contract at a price of S_t is thus more expensive than buying the dollars immediately for the same price since in the former case the investor can invest the money up to time T in a domestic bond paying an interest rate which is higher than the American interest rate.

The following result states that forward and future contracts with the same delivery price and the same time to maturity are equal, if interest rates are constant during the contract period. We will use the fact that by definition forward and future contracts do not cost anything if the delivery price is chosen to be equal to the current price of the forward contract respectively the price of the future contract.

Theorem 2.2
If interest rates are constant during contract period, then forward and future prices are equal.

Proof:
We proceed from the assumption that the future contract is agreed on at time 0, and that is has a time to maturity of N days. We assume that profits and losses are settled (marked to market) on a daily basis at a daily interest rate of ρ. While the forward price at the end of day 0 is denoted by F, the future price at the end of day t, $t = 0, 1, \cdots, N$ is denoted by F_t. The goal is to show that $F = F_0$. For that we construct two portfolios again:

Portfolio A**:** A long position in $e^{N\rho}$ forward contracts with delivery price F and maturity date N.
A long position in a zero bond with face value $Fe^{N\rho}$ maturing in N days.

Portfolio B**:** A long position in futures contracts with delivery price F_t and maturity date N. The contracts are bought daily such that the portfolio contains at the end of the t–th day exactly $e^{(t+1)\rho}$ future contracts $(t = 0, 1, \cdots, N)$.
A long position in a zero bond with face value $F_0e^{N\rho}$ maturing in N days.

Purchasing a forward or a future contract does not cost anything since their delivery prices are set to equal the current forward or future price. Due to the marking to market procedure the holder of portfolio B receives from day $t-1$ to day t for each future contract an amount of $F_t - F_{t-1}$ which can possibly be negative (i.e. he has to pay).

At maturity, i.e. at the end of day N, the zero bond of portfolio A is sold at the face value $Fe^{N\rho}$ to fulfill the terms of the forward contract and to buy $e^{N\rho}$ stocks at a the delivery price F. Then A contains exclusively these stocks and has a value of $S_N e^{N\rho}$. Following, we show that portfolio B has the same value.

At the beginning of day t portfolio B contains $e^{t\rho}$ future contracts, and the holder receives due to the marking to market procedure the amount $(F_t - F_{t-1})e^{t\rho}$ which can possibly be negative. During the day he increases his long position in the future contracts at zero costs such that the portfolio contains $e^{(t+1)\rho}$ future contracts at the end of the day. The earnings at day t compounded to the maturity date have a value of:

$$(F_t - F_{t-1})e^{t\rho} \cdot e^{(N-t)\rho} = (F_t - F_{t-1})e^{N\rho}. \tag{2.10}$$

At maturity the terms of the future contracts are fulfilled due to the marking to market procedure. All profits and losses compounded to day N have a value of:

$$\sum_{t=1}^{N}(F_t - F_{t-1})e^{N\rho} = (F_N - F_0)e^{N\rho}. \tag{2.11}$$

Together with the zero bond portfolio B has at day N a value of

$$(F_N - F_0)e^{N\rho} + F_0 e^{N\rho} = F_N e^{N\rho} = S_N e^{N\rho},$$

since at maturity the future price F_N and the price S_N of the underlying are obviously equal.

Hence, both portfolios have at day N the same value and thus due to the no–arbitrage assumption their day 0 values must be equal as well. Since the forward contract with delivery price F has a value of 0 at day 0 due to the definition of the forward price, the value of portfolio A is equal to the value of the zero bond, i.e. F (face value $Fe^{N\rho}$ discounted to day 0). Correspondingly, the e^{ρ} futures contained in portfolio B have at the end of day 0 a value of 0 due to the definition of the future price. Again, the value of portfolio B reduces to the value of the zero bond. The latter has a value of F_0 (face value $F_0 e^{N\rho}$ discounted to day 0). Putting things together, we conclude that $F = F_0$. □

Now, we want to proof some relationship between option prices using similar methods. The most elementary properties are summarized in the following remark without a proof. For that, we need the notion of the *intrinsic value* of an option.

Definition 2.1 (Intrinsic Value)
The intrinsic value *of a call option at time t is given by* $\max(S_t - K, 0)$, *the intrinsic value of a put option is given by* $\max(K - S_t, 0)$. *If the intrinsic value of an option is positive we say that the option is* in the money. *If $S_t = K$, then the option is* at the money. *If the intrinsic value is negative, then the option is said to be* out of the money.

Remark 2.1
Options satisfy the following elementary relations. $C(s,\tau) = C_{K,T}(s,\tau)$ and $P(s,\tau) = P_{K,T}(s,\tau)$ denote the time t value of a call and a put with delivery price K and maturity date T, if $\tau = T - t$ is the time to maturity and the price of the underlying is s, i.e. $S_t = s$.

1. *Option prices are non negative since an exercise only takes place if it is in the interest of the holder. An option gives the right to exercise. The holder is not obligated to do so.*
2. *American and European options have the same value at maturity T since in T they give the same rights to the holder. At maturity T the value of the option is equal to the intrinsic value:*

$$C_{K,T}(S_T, 0) = \max(S_T - K, 0)\ , \quad P_{K,T}(S_T, 0) = \max(K - S_T, 0).$$

3. *An American option must be traded at least at its intrinsic value since otherwise a riskless profit can be realized by buying and immediately exercising the option. This relation does not hold in general for European options. The reason is that a European option can be exercised only indirectly by means of a future contract. The thereby involved discounting rate can possibly lead to the option being worth less than its intrinsic value.*
4. *The value of two American options which have different time to maturities, $T_1 \le T_2$, is monotonous in time to maturity:*

$$C_{K,T_1}(s, T_1 - t) \le C_{K,T_2}(s, T_2 - t)\ , \quad P_{K,T_1}(s, T_1 - t) \le P_{K,T_2}(s, T_2 - t).$$

This follows, for calls, say, using 2., 3. from the inequality which holds at time $t = T_1$ with $s = S_{T_1}$

$$C_{K,T_2}(s, T_2 - T_1) \ge \text{ intrinsic value } = \max(s - K, 0) = C_{K,T_1}(s, 0) \tag{2.12}$$

Due to the no–arbitrage assumption the inequality must hold for any point in time $t \le T_1$. For European options this result does not hold in general.

5. *An American option is at least as valuable as the identically specified European option since the American option gives more rights to the holder.*

6. *The value of a call is a monotonously decreasing function of the delivery price since the right to buy is the more valuable the lower the agreed upon delivery price. Accordingly, the value of a put is a monotonously increasing function of the delivery price.*

$$C_{K_1,T}(s,\tau) \geq C_{K_2,T}(s,\tau)\ , \quad P_{K_1,T}(s,\tau) \leq P_{K_2,T}(s,\tau)$$

   for $K_1 \leq K_2$. *This holds for American as well as for European options.*

The value of European call and put options on the same underlying with the same time to maturity and delivery price are closely linked to each other without using a complicated mathematical model.

Theorem 2.3 (Put–Call Parity for European Options)
For the value of a European call and put option which have the same maturity date T, *the same delivery price* K, *the same underlying the following holds (where* r *denotes the continuous interest rate):*

1. *If the underlying pays a dividend yield with a time* t *discounted total value of* D_t *during the time to maturity* $\tau = T - t$ *then it holds*

$$C(S_t,\tau) = P(S_t,\tau) + S_t - D_t - Ke^{-r\tau} \tag{2.13}$$

`SFEPutCall.xpl`

2. *If the underlying involves continuous costs of carry at rate* b *during the time to maturity* $\tau = T - t$ *then it holds*

$$C(S_t,\tau) = P(S_t,\tau) + S_t e^{(b-r)\tau} - Ke^{-r\tau} \tag{2.14}$$

Proof:
For simplicity, we again assume the underlying to be a stock. We consider a portfolio A consisting of one call which will be duplicated by a suitable portfolio B containing a put among others.

1\. In the case of discrete dividend yields we consider at time t the following portfolio B :

1. Buy the put.

	Value at timet T	
Position	$K < S_T$	$K \geq S_T$
a)	0	$K - S_T$
b)	$-K$	$-K$
c)	S_T	S_T
d)	0	0
Sum	$S_T - K$	0

Table 2.1: Value of portfolio B at time T (Theorem 2.3).

2. Sell a zero bond with face value K maturing T.
3. Buy one stock.
4. Sell a zero bond at the current price D_t.

The stock in portfolio B pays dividends whose value discounted to time t is D_t. At time T these dividend yields are used to pay back the zero bond of position d). Hence this position has a value of zero at time T. Table 2.1 shows the value of portfolio B at time T where we distinguished the situations where the put is exercised ($K \geq S_T$) and where it is not exercised. At time T portfolio B has thus the same value $\max(S_T - K, 0)$ as the call. To avoid arbitrage opportunities both portfolios A and B must have the same value at any time t prior T, that is it holds

$$C(S_t, \tau) = P(S_t, \tau) - Ke^{-r\tau} + S_t - D_t \tag{2.15}$$

2. In the case of continuous dividends at rate d and corresponding costs of carry $b = r - d$ we consider the same portfolio B as in part 1. but this time without position d). Instead we buy $e^{-d\tau}$ stocks in position c) whose dividends are immediately reinvested in the same stock. If d is negative, then the costs are financed by selling stocks. Thus, portfolio B contains exactly one stock at time T, and we conclude as in part 1. that the value of portfolio B is at time t equal to the value of the call. □

The proof of the put–call parity holds only for European options. For American options it may happen that the put or call are exercised prior maturity and that both portfolios are not hold until maturity.

The following result makes it possible to check whether prices of options on the same underlying are consistent. If the convexity formulated below is

	Value at time t'			
Position	$S_{t'} \leq K_1$	$K_1 < S_{t'} \leq K_\lambda$	$K_\lambda < S_{t'} \leq K_0$	$K_0 < S_{t'}$
B 1.	0	$\lambda(S_{t'} - K_1)$	$\lambda(S_{t'} - K_1)$	$\lambda(S_{t'} - K_1)$
B 2.	0	0	0	$(1-\lambda)(S_{t'} - K_0)$
$-A$	0	0	$-(S_{t'} - K_\lambda)$	$-(S_{t'} - K_\lambda)$
Sum	0	$\lambda(S_{t'} - K_1)$	$(1-\lambda)(K_0 - S_{t'})$	0

Table 2.2: Difference in the values of portfolios B and A at time t' (Theorem 2.4).

violated, then arbitrage opportunities arise as we will show in the example following the proof of the next theorem.

Theorem 2.4
The price of a (American or European) Option is a convex function of the delivery price.

Proof:
It suffices to consider calls since the proof is analogous for puts. The put–call parity for European options is linear in the term which depends explicitly on K. Hence, for European options it follows immediately that puts are convex in K given that calls are convex in K.

For $0 \leq \lambda \leq 1$ and $K_1 < K_0$ we define $K_\lambda \stackrel{\text{def}}{=} \lambda K_1 + (1-\lambda)K_0$. We consider a portfolio A which at time $t < T$ consists of one call with delivery price K_λ and maturity date T. At time t we duplicate this portfolio by the following portfolio B :

1. A long position in λ calls with delivery price K_1 maturing in T.
2. A long position in $(1-\lambda)$ calls delivery price K_0 maturing in T.

By liquidating both portfolios at an arbitrary point of time $t', t \leq t' \leq T$ we can compute the difference in the values of portfolio A and B which is given in Table 2.2

Since $\lambda(S_{t'} - K_1) \geq 0$ und $(1-\lambda)(K_0 - S_{t'})) \geq 0$ in the last row of Table 2.2 the difference in the values of portfolio A and B at time t' and thus for any point of time $t < t'$ is greater than or equal to zero. Hence, denoting $\tau = T - t$ it holds

$$\lambda C_{K_1,T}(S_t,\tau) + (1-\lambda)C_{K_0,T}(S_t,\tau) - C_{K_\lambda,T}(S_t,\tau) \geq 0 \qquad (2.16)$$

Delivery price	Option price
$K_1 = 190$	30.6 EUR
$K_\lambda = 200$	26.0 EUR
$K_0 = 220$	14.4 EUR

Table 2.3: Data of Example 2.3.

□

Example 2.3
*We consider three European call options on the MD*TECH A.G. having all the same time to maturity and delivery prices* $K_1 = 190$, $K_\lambda = 200$, $K_0 = 220$, *i.e.* $\lambda = \frac{2}{3}$. *Table 2.3 shows the data of this example. Due to the last theorem it must hold:*

$$\tfrac{2}{3} C_{K_1,T}(S_t, \tau) + \tfrac{1}{3} C_{K_0,T}(S_t, \tau) \geq C_{K_\lambda,T}(S_t, \tau) \tag{2.17}$$

Since this condition is obviously violated an arbitrage opportunity exists, and with the following portfolio a riskless gain can be realized:

1. *A long position in* $\lambda = \frac{2}{3}$ *calls with delivery price* K_1.
2. *A long position in* $1 - \lambda = \frac{1}{3}$ *calls with delivery price* K_0.
3. *A short position in 1 call with delivery price* $K_\lambda \stackrel{\text{def}}{=} \frac{2}{3} K_1 + \frac{1}{3} K_0$.

By setting up this portfolio at the current time t *we realize an immediate profit of +0.80* EUR. *The portfolio value at options' maturity* T *is given by Table 2.4 from which we can extract that we realize further profits for stock prices* S_T *between 190 and 220 of at most* $\dfrac{20}{3}$ EUR.

We already said that option prices are monotonous functions of the delivery price. The following theorem for European options is more precise on this subject.

Theorem 2.5
For two European calls (puts) with the same maturity date T *and delivery prices* $K_1 \leq K_2$ *it holds at time* $t \leq T$:

$$0 \leq C_{K_1,T}(S_t, \tau) - C_{K_2,T}(S_t, \tau) \leq (K_2 - K_1) e^{-r\tau} \tag{2.18}$$

	Value at time T			
Position	$S_T \leq 190$	$190 < S_T \leq 200$	$200 < S_T \leq 220$	$220 < S_T$
1.	0	$\frac{2}{3}(S_T - 190)$	$\frac{2}{3}(S_T - 190)$	$\frac{2}{3}(S_T - 190)$
2.	0	0	0	$\frac{1}{3}(S_T - 220)$
3.	0	0	$-(S_T - 200)$	$-(S_T - 200)$
Sum	0	$\frac{2}{3}(S_T - 190)$	$\frac{1}{3}(220 - S_T)$	0

Table 2.4: Portfolio value at time T of Example 2.3.

or

$$0 \leq P_{K_2,T}(S_t,\tau) - P_{K_1,T}(S_t,\tau) \leq (K_2 - K_1)e^{-r\tau} \tag{2.19}$$

with $\tau = T - t$ and r denoting the time to maturity and the interest rate respectively. If call (put) option prices are differentiable as a function of the delivery price, then by taking the limit $K_2 - K_1 \rightarrow 0$ it follows

$$1 \leq -e^{-r\tau} \leq \frac{\partial C}{\partial K} \leq 0 \quad \textit{bzw.} \quad 0 \leq \frac{\partial P}{\partial K} \leq e^{-r\tau} \leq 1 \tag{2.20}$$

Proof:
We proof the theorem for calls since for puts the reasoning is analogous. For this we consider a portfolio A containing one call with delivery price K_1 which we compare to a duplicating portfolio B. At time t the latter portfolio consists of the following two positions:

1. A long position in one call with delivery price K_2.
2. A long position in one zero bond with face value $(K_2 - K_1)$ maturing in T.

The difference of the value of portfolios B and A at time T is shown in Table 2.5. At time T portfolio B is clearly as valuable as portfolio A which given the no–arbitrage assumption must hold at time t as well. We conclude:

$$C_{K_2,T}(S_t,\tau) + (K_2 - K_1)e^{-r\tau} \geq C_{K_1,T}(S_t,\tau).$$

□

2.2 Portfolio Insurance

A major purpose of options is hedging, i.e. the protection of investments against market risk caused by random price movements. An example for

Position	Value at time T $S_T \leq K_1$	$K < S_T < K_2$	$K_2 \leq S_T$
B 1.	0	0	$S_T - K_2$
B 2.	$K_2 - K_1$	$K_2 - K_1$	$K_2 - K_1$
$-A$	0	$-(S_T - K_1)$	$-(S_T - K_1)$
Sum	$K_2 - K_1$	$K_2 - S_T$	0

Table 2.5: Difference in the values of portfolios B and A at time T (Theorem 2.5).

active hedging with options is the portfolio insurance. That is to strike deals in order to change at a certain point of time the risk structure of a portfolio such that at a future point of time

- the positive profits are reduced by a small amount (which can be interpreted as an insurance premium) and in that way
- the portfolio value does not drop below a certain *floor*.

The portfolio insurance creates a risk structure of the portfolio which prevents extreme losses. For illustration purposes we consider at first a simple example.

Example 2.4
An investor has a capital of 10 500 EUR *at his disposal to buy stocks whose current price is 100* EUR*. Furthermore, put options on the same stock with a delivery price of* $K = 100$ *and a time to maturity of one year are quoted at a market price of 5* EUR *per contract. We consider two investment alternatives.*

Portfolio A: *Buying 105 stocks.*

Portfolio B: *Buying 100 stocks for 10 000* EUR *and buying 100 put options for 500* EUR*.*

The price of the put options can be interpreted as the premium to insure the stocks against falling below a level of 10 000 EUR*. Denoting the stock price in one year by* S_T *the value of the non–insured portfolio is* $105 \cdot S_T$*. This portfolio bears the full market risk that the stock price drops significantly below 100* EUR*. The insured portfolio, however, is at least as worth as 10 000* EUR

Stock price S_T [EUR]	Non–insured portfolio Value [EUR]	Non–insured portfolio Return % p.a.	Insured portfolio Value [EUR]	Insured portfolio Return % p.a.	Insured portfolio in % of the non–insured portfolio
50	5250	–50	10000	–4.8	190
60	6300	–40	10000	–4.8	159
70	7350	–30	10000	–4.8	136
80	8400	–20	10000	–4.8	119
90	9450	–10	10000	–4.8	106
100	10500	0	10000	-4.8	95
110	11550	+10	11000	+4.8	95
120	12600	+20	12000	+14.3	95
130	13650	+30	13000	+23.8	95
140	14700	+40	14000	+33.3	95

Table 2.6: The effect of a portfolio insurance on portfolio value and return.

since if $S_T < 100$ *the holder exercises the put options and sells the 100 stocks for 100* EUR *each.*

Should the stock price increase above 100 EUR *the investor does not exercise the put which thus expires worthless. By buying the put some of the capital of portfolio B is sacrificed to insure against high losses. But, while the probabilities of high profits slightly decrease, the probabilities of high losses decrease to zero. Investing in portfolio B the investor looses at most 500* EUR *which he paid for the put. Table 2.6 shows the impact of the stock price* S_T *in one year on both the insured and the non–insured portfolio values and returns.*

The numerous conceivable strategies to insure portfolios can be classified by the frequency with which the positions in the portfolio have to rebalanced. Two approaches can be distinguished:

- Static strategies rebalance the portfolio positions at most once before expiration of the investment horizon.
- Dynamic strategies rebalance the portfolio positions very frequently, ideally continuously, according to certain rules.

The static strategy sketched in the previous example can be modified. Instead of hedging by means of put options the investor can chose between the following two strategies:

Strategy 1: The investor buys an equal number of *stocks and puts.*

Strategy 2: The investor buys *bonds* with a face value equal to the floor he is aiming at and for the remaining money he buys calls on the stock.

All strategies commonly practiced rely on modifications of the above basic strategies. While following the first strategy it is the put which guarantees that the invested capital does not drop below the floor, applying the second strategy it is the bond which insures the investor against falling prices. The stocks respectively the calls make up for the profits in case of rising prices. The equivalence of both strategies follows from the put–call parity, see Theorem 2.3.

Before deciding about what kind of portfolio insurance will be used some points have to be clarified:

1. Which financial instruments are provided by the market, and what are their characteristics (coupons, volatilities, correlation with the market etc.)?
2. Which ideas does the investor have about
   - the composition of the portfolio (which financial instruments),
   - the amount of capital to invest,
   - the investment horizon,
   - the floor (lower bound of the portfolio value) or rather the minimum return he is aiming at the end of the investment. Given the floor F and the capital invested V the possibly negative minimum return of a one year investment is given by $\rho = \frac{F-V}{V}$.

The strategies 1 and 2 described above we illustrate in another example.

Example 2.5
We proceed from the assumption that the investor has decided to invest in stock. Depending on the type of return of the object we distinguish two cases (for negative returns, as storage costs of real values for example, the approach can be applied analogously):

i) *continuous dividend yield d*

ii) *ex ante known discrete yields with a time 0 discounted total value of D_0.*

Data of Example 2.5:	
Current point of time t	0
Available capital V	100 000 EUR
Target floor F	95 000 EUR
Investment horizon T	2 years
Current stock price S_0	100 EUR
Continuously compounded annual interest rate r	0.10
Annual stock volatility σ	0.30
Dividends during time to maturity	
Case i): continuous dividends d	0.02
Fall ii): dividends with present value D_0	5 EUR

Table 2.7: Data of Example 2.5

The data of the example is shown in Table 2.7. The volatility can be interpreted as a measure of variability of the stock price. The notion of volatility is an essential part of option pricing and will be treated extensively later. Placing our considerations at the beginning $t = 0$ of the investment the time to maturity is $\tau = T - t = T$. For both strategies the goal is to determine the number n of stocks and/or (European) options and their delivery price K.

Case i)
The stock pays a continuous dividend at rate $d = 2\%$ p.a. which he reinvests immediately. At maturity T the position in the stock grew from n stocks to $ne^{d\tau}$ with $\tau = T - 0 = T$. Thus, for strategy 1 he has to buy in $t = 0$ the same number of put options. Since the amount he wants to invest in $t = 0$ is V it must hold

$$n \cdot S_0 + ne^{d\tau} \cdot P_{K,T}(S_0, \tau) = V. \tag{2.21}$$

The investor chooses the put options delivery price K such that his capital after two years does not drop below the floor F he is aiming at. That is, exercising the puts in time T (if $S_T \le K$) must give the floor F which gives the second condition

$$ne^{d\tau} \cdot K = F \iff n = \frac{F}{K} e^{-d\tau}. \tag{2.22}$$

Substituting equation (2.22) into equation (2.21) gives

$$e^{-d\tau} S_0 + P_{K,T}(S_0, \tau) - \frac{V}{F} \cdot K = 0. \tag{2.23}$$

Thanks to the Black–Scholes pricing formula for European options that will be derived later in Section 6.2 the put price is expressed as a function of

Stock price S_T [EUR]	Non–insured portfolio Value [EUR]	Non–insured portfolio Return % p.a.	Insured portfolio Value [EUR]	Insured portfolio Return % p.a.	Insured portfolio in % of the non–insured portfolio
70	72857	–27	95000	–5	130
80	83265	–17	95000	–5	114
90	93673	–6	95000	–5	101
100	104081	+4	95400	–5	92
110	114489	+15	104940	+5	92
120	124897	+25	114480	+14	92
130	135305	+35	124020	+24	92
140	145714	+46	133560	+34	92

Table 2.8: The effect of a portfolio insurance in case i) on portfolio value and return. `SFEoptman.xpl`

the parameter K. The delivery price which he is looking for can be computed by solving equation (2.23) numerically, for example by means of the Newton–Raphson method. In this case, K is equal to 99.56. *To be sure that the capital value does not drop below the floor $F = 95\ 000$ EUR he buys $n = \frac{F}{K}e^{-d\tau} = 916.6$ stocks and $n \cdot e^{d\tau} = 954$ puts with delivery price $K = 99.56$. The price of the put option given by the Black–Scholes formula is* 8.72 EUR*/put.* `SFEexerput.xpl`

Following the corresponding strategy 2 he invests $Fe^{-r\tau} = 77\,779.42$ EUR in bonds at time 0 which gives compounded to time T exactly the floor $F = 95\ 000$ EUR. For the remaining capital of $V - Fe^{-r\tau} = 22\ 220.58$ EUR he buys 954 *calls with delivery price $K = 99.56$ which have a price of* 23.29 EUR*/call according to the Black–Scholes formula. From the put–call parity follows the equivalence of both strategies, i.e. both portfolios consisting of stocks and puts respectively zero bonds and calls have at each time t the same value:*

$$n \cdot S_t + ne^{d\tau} P_{K,T}(S_t, \tau) = nKe^{-b\tau} + ne^{d\tau} C_{K,T}(S_0, \tau) \tag{2.24}$$

where $\tau = T - t, b = r - d$ and $nKe^{-b\tau} = Fe^{-r\tau}$ due to equation (2.22). Table 2.8 shows the risk decreasing effect of the insurance.

Case ii)

Until maturity the stock pays dividends with a time 0 discounted total value $D_0 = 5$ EUR which are after distribution immediately invested in bonds. At time T the dividend yield has a compounded value of $D_T = D_0 e^{r\tau} = 6.107$

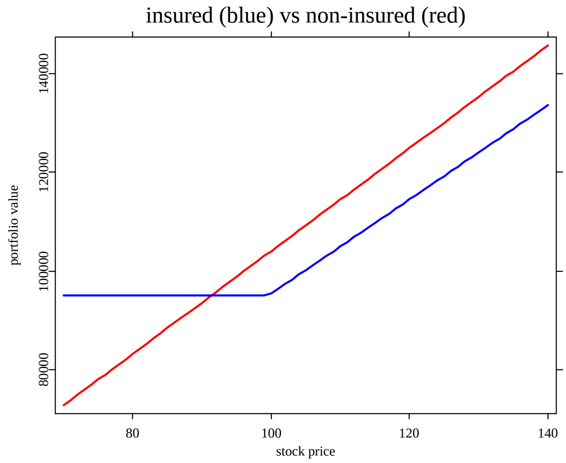


Figure 2.1: The effect of a portfolio insurance: While the straight line represents the value of the insured portfolio as a function of the stock price, the dotted line represents the value of the non–insured portfolio as a function of the stock price. SFEoptman.xpl

EUR *where* $\tau = T$ *denotes the time to maturity. Reasoning as in case i) and taking the dividend* D_T *into account he buys* n *stocks respectively* n *puts, and obtains the following equations*

$$n \cdot S_0 + nP_{K,T}(S_0 - D_0, \tau) = V \tag{2.25}$$

and

$$nK + nD_T = F. \tag{2.26}$$

The substraction of the cash dividend D_0 *from the stock price* S_0 *in the option price formula cannot be justified until we introduced the binomial model in Chapter 7. Briefly, in a perfect market the stock price decreases instantaneously by the amount of the distributed dividend. Otherwise, an arbitrage opportunity arises. Substituting equation (2.26) into equation (2.25) gives:*

$$S_0 + P_{K,T}(S_0 - D_0, \tau) - \frac{V}{F} \cdot (K + D_T) = 0 \tag{2.27}$$

Solving the equations analogously as in case i) the number n *of stocks and puts and the delivery price* K *for strategy 1 are obtained:*

$$K = 96.42 \quad \textit{und} \quad n = \frac{F}{K + D_T} = 926.55$$

Stock price S_T [EUR]	Non–insured portfolio Value [EUR]	Non–insured portfolio Return % p.a.	Insured portfolio Value [EUR]	Insured portfolio Return % p.a.	Insured portfolio in % of the non–insured portfolio
70	76107	–24	94996	–5	125
80	86107	–14	94996	–5	110
90	96107	–4	94996	–5	99
96.42	102527	+3	94996	–5	93
100	106107	+6	98313	–2	93
110	116107	+16	107579	+8	93
120	126107	+26	116844	+17	93
130	136107	+36	126110	+26	93
140	146107	+46	135375	+35	93

Table 2.9: The effect of a portfolio insurance in case ii) on portfolio value and return.

For strategy 2 he buys 926.55 *calls at a price of* 23.99 EUR*/call with a delivery price* $K = 96.42$. *He invests* $95\,000e^{-r\tau} = 77\,779.42$ *in bonds. For case ii) the effect of the portfolio insurance for both strategies is shown in Table 2.9 taking into account the time* T *compounded total dividend.*

The example shows how a portfolio insurance can be carried out by means of options in principle. In practice, the following problems frequently occur:

- The number n of stocks and options is not an integer. In a perfect financial market financial instruments are perfectly divisible, in reality, however, this is not the case. The error resulting from rounding up or down to closest integer can be neglected only in large portfolios.
- Puts and calls traded on the market do not cover the whole range of delivery prices. Thus, options with the computed delivery price are possibly not available. Furthermore, options typically expire in less than one year which makes static strategies only limited applicable when long–term investments are involved.
- Finally, the market provides first of all American options which are more expensive than European options which are sufficient to insure the portfolio. The additional exercise opportunities offered by American options, are only of interest if the investor possibly has to close the portfolio early.

The fact that options are not traded at all delivery prices suggests to produce them by the delta hedge process described in Chapter 6. But since a dynamic strategy is involved transaction costs have to be taken into account and give rise to other problems. Finally, we point out that when insuring large portfolios it is convenient to hedge by means of index options, i.e. puts and calls on the DAX for example, not only from a cost saving point of view but also those options replace options on a single underlying which are not traded on the market. To compute the exact effect of an index option hedge the correlation of the portfolio with the index is needed. The latter correlation is obtained from the correlations of each individual stock contained in the portfolio with the index. Besides detailed model assumptions as the Capital Asset Pricing Model (CAPM see Section 10.4.1) which among others concern the stock returns are required.

2.3 Recommended Literature

The fundamental arbitrage relations of financial derivatives can be found in every modern finance textbook, as for example Hull (2000). In principle each option pricing theory is based on theses relations, as it is the case for the model of Black and Scholes (1973) for example, see also the fundamental article of Merton (1973). The idea of portfolio assurance was introduced by Leland (1980). It is nowadays covered above all by practical risk management textbooks as Das (1997).

3 Basic Concepts of Probability Theory

3.1 Real Valued Random Variables

Thanks to Newton's laws, dropping a stone from a latitude of 10 m, the point of time of its impact on the ground is known before executing the experiment. Quantities in complex systems (such as stock prices at a certain date, daily maximum temperature at a certain place) are however not deterministically predictable, although it is known which values are more likely to occur than others. Contrary to the falling stone, data which cannot be described successfully by deterministic mechanism, can be modelled by random variables.

Let X be such a random variable which, (as a model for stock prices), takes values on the real time. The appraisal of which values of X are more and which are less likely, is expressed by the *probability* of events as $\{a < X < b\}$ or $\{X \le b\}$. The set of all probabilities

$$\mathrm{P}(a \le X \le b)\ , \quad -\infty < a \le b < \infty\ ,$$

determines the *distribution* of X. In other words, the distribution is defined by the probabilities of all events which depend on X. In the following, we denote the probability distribution of X by $\mathcal{L}(X)$.

The probability distribution is uniquely defined by the *cumulative probability distribution*

$$F(x) = \mathrm{P}(X \le x)\ , \quad -\infty < x < \infty\ .$$

$F(x)$ is monotonously increasing and converges for $x \to -\infty$ to 0, and for $x \to \infty$ to 1. If there is a function p, such that the probabilities can be computed by means of an integral

$$\mathrm{P}(a < X < b) = \int_a^b p(x)dx\,,$$

p is called a *probability density*, or briefly density of X. Then the cumulative distribution function is a primitive of p:

$$F(x) = \int_{-\infty}^{x} p(y)dy.$$

For small h it holds:

$$\mathrm{P}(x - h < X < x + h) \approx 2h \cdot p(x).$$

Thus $p(x)$ is a measure of the likelihood that X takes values close to x.

The most important family of distributions with densities, is the *normal distribution* family. It is characterized by two parameters μ, σ^2. The densities are given by

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma}\,\varphi\big(\frac{x-\mu}{\sigma}\big)\,,$$

$$\varphi(x) = \varphi_{0,1}(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\,.$$

The distribution with density $\varphi(x)$ is called *standard normal distribution.* "X is a normal random variable with parameters μ, σ^2" is commonly abbreviated by "X is $N(\mu, \sigma^2)$ distributed". The cumulative distribution function of a standard normal distribution is denoted by Φ and it holds:

$$\Phi(x) = \int_{-\infty}^{x} \varphi(y)dy.$$

If X is $N(\mu, \sigma^2)$ distributed, then the centered and scaled random variable $(X - \mu)/\sigma$ is standard normal distributed. Therefore, its cumulative distribution function is given by:

$$F(x) = \mathrm{P}(X \le x) = \mathrm{P}\big(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\big) = \Phi\big(\frac{x-\mu}{\sigma}\big).$$

A distribution which is of importance in modelling stock prices is the *lognormal distribution.* Let X be a positive random variable whose natural logarithm $\ln(X)$ is normally distributed with parameters μ, σ^2. We say that X is lognormally distributed with parameters μ, σ^2. Its cumulative distribution function follows directly from the above definition:

$$F(x) = \mathrm{P}(X \le x) = \mathrm{P}(\ln X \le \ln x) = \Phi\big(\frac{\ln x - \mu}{\sigma}\big), \quad x > 0.$$

Deriving $F(x)$ once, we obtain its density function with parameters μ, σ^2:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} = \frac{1}{\sigma x} \varphi\Big(\frac{\ln x - \mu}{\sigma}\Big), \quad x > 0.$$

If X is a random variable that takes only finitely or countably infinite values $x_1, \ldots, x_n$, X is said to be a discrete random variable and its distribution is fully determined by the probabilities:

$$\mathrm{P}(X = x_k) \ , \quad k = 1, \ldots, n.$$

The simplest discrete random variables take only 2 or 3 values, for example ± 1 or $-1, 0, +1$. They constitute the basis of binomial or trinomial trees which can be used to construct discrete random processes in computers. Such tree methods are reasonable approximations to continuous processes which are used to model stock prices.

In this context, *binomially distributed* random variables appear. Let $Y_1, \ldots, Y_n$ be independent random variables taking two values, 0 or 1, with probabilities

$$p = \mathrm{P}(Y_k = 1) \ , \ 1 - p = \mathrm{P}(Y_k = 0) \ , \quad k = 1, \ldots, n \, .$$

We call such random variables *Bernoulli distributed* with parameter p. The number of ones appearing in the sample $Y_1, \ldots, Y_n$, equals the sum $X = \sum_{k=1}^n Y_k$ which is binomial distributed with parameters n, p :

$$X = \sum_{k=1}^{n} Y_k \ , \ \mathrm{P}(X = m) = \binom{n}{m} p^m (1-p)^{n-m} \ , \quad m = 0, \ldots, n \, .$$

SFEBinomial.xpl

Instead of saying X is binomial distributed with parameters n, p, we use the notation "X is $B(n, p)$ distributed". Hence, a Bernoulli distributed random variable is $B(1, p)$ distributed.

If n is large enough, a $B(n, p)$ distributed random variable can be approximated by a $N(np, \, np(1-p))$ distributed random variable Z, in the sense that

$$\mathrm{P}(a < X < b) \approx \mathrm{P}(a < Z < b) \, . \tag{3.1}$$

The central limit theorem is more precise on that matter. In classical statistics it is used to avoid, for large n, the tedious calculation of binomial probabilities. Conversely, it is possible to approximate the normal distribution by an easy simulated binomial tree. SFEclt.xpl

3.2 Expectation and Variance

The mathematical *expectation* or the *mean* $\mathsf{E}[X]$ of a real random variable X is a measure for the location of the distribution of X. Adding to X a real constant c, it holds for the expectation: $\mathsf{E}[X+c] = \mathsf{E}[X]+c$, i.e. the location of the distribution is translated. If X has a density $p(x)$, its expectation is defined as:

$$\mathsf{E}(X) = \int_{-\infty}^{\infty} xp(x)dx.$$

If the integral does not exist, neither does the expectation. In practice, this is rather rarely the case.

Let $X_1, \ldots, X_n$ be a sample of identically independently distributed (i.i.d.) random variables (see Section 3.4) having the same distribution as X, then $\mathsf{E}[X]$ can be estimated by means of the sample mean:

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} X_t.$$

A measure for the dispersion of a random variable X around its mean is given by the *variance* $\mathrm{Var}(X)$:

$$\begin{aligned} \mathrm{Var}(X) &= \mathsf{E}[(X - \mathsf{E}\,X)^2] \\ \text{Variance} &= \text{mean squarred deviation of a random variable} \\ & \text{around its expectation.} \end{aligned}$$

If X has a density $p(x)$, its variance can be computed as follows:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - \mathsf{E}\,X)^2 p(x)dx.$$

The integral can be infinite. There are empirical studies giving rise to doubt that some random variables appearing in financial and actuarial mathematics and which model losses in highly risky businesses dispose of a finite variance.

As a quadratic quantity the variance's unity is different from that of X itself. It is better to use the standard deviation of X which is measured in the same unity as X:

$$\sigma(X) = \sqrt{\mathrm{Var}(X)}\,.$$

Given a sample of i.i.d. variables $X_1, \ldots, X_n$ which have the same distribution as X, the sample variance can be estimated by:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} (X_t - \hat{\mu})^2.$$

A $N(\mu, \sigma^2)$ distributed random variable X has mean μ and variance σ^2. The 2σ area around μ contains with probability of a little more than 95% observations of X:

$$\mathrm{P}(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.95\,.$$

A lognormally distributed random variable X with parameters μ and σ^2 has mean and variance

$$\mathsf{E}(X) = e^{\mu + \frac{1}{2}\sigma^2}, \quad \mathrm{Var}(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1).$$

A $B(n, p)$ distributed variable X has mean np and variance $np(1 - p)$. The approximation (3.1) is chosen such that the binomial and normal distribution have identical mean and variance.

3.3 Skewness and Kurtosis

Definition 3.1 (Skewness)
The skewness of a random variable X with mean μ and variance σ^2 is defined as

$$S(X) = \frac{\mathsf{E}[(X - \mu)^3]}{\sigma^3}.$$

If the skewness is negative (positive) the distribution is skewed to the left (right). Normally distributed random variables have a skewness of zero since the distribution is symmetrical around the mean. Given a sample of i.i.d. variables $X_1, \ldots, X_n$, Skewness can be estimated by (see Section 3.4)

$$\hat{S}(X) = \frac{\frac{1}{n}\sum_{t=1}^{n}(X_t - \hat{\mu})^3}{\hat{\sigma}^3}, \tag{3.2}$$

with $\hat{\mu}, \hat{\sigma}^2$ as defined in the previous section.

Definition 3.2 (Kurtosis)
The kurtosis of a random variable X with mean μ and variance σ^2 is defined as

$$\mathrm{Kurt}(X) = \frac{\mathsf{E}[(X - \mu)^4]}{\sigma^4}.$$

Normally distributed random variables have a kurtosis of 3. Financial data often exhibits higher kurtosis values, indicating that values close to the mean and extreme positive and negative outliers appear more frequently than for normally distributed random variables. Kurtosis can be estimated by

$$\widehat{\mathrm{Kurt}}(X) = \frac{\frac{1}{n}\sum_{t=1}^{n}(X_t-\hat{\mu})^4}{\hat{\sigma}^4}. \tag{3.3}$$

Example 3.1 *The empirical standard deviation of monthly DAX data from 1979:1 to 2000:10 is* $\hat{\sigma} = 0.056$*, which corresponds to a yearly volatility of* $\hat{\sigma} \cdot \sqrt{12} = 0.195$*. Later in Section(6.3.4), we will explain the factor* $\sqrt{12}$ *in detail. The kurtosis of the data is much greater than 3 which suggests a non-normality of the DAX returns.*

SFEsumm.xpl

3.4 Random Vectors, Dependence, Correlation

A *random vector* $(X_1, \ldots, X_n)$ from $\mathbb{R}^n$ can be useful in describing the mutual dependencies of several random variables $X_1, \ldots, X_n$, for example several underlying stocks. The joint distribution of the random variables $X_1, \ldots, X_n$ is as in the univariate case, uniquely determined by the probabilities

$$\mathrm{P}(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n)\ , \quad -\infty < a_i \le b_i < \infty\,,\ i = 1, ..., n\,.$$

If the random vector $(X_1, \ldots, X_n)$ has a density $p(x_1, \ldots, x_n)$, the probabilities can be computed by means of the following integrals:

$$\mathrm{P}(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n) = \int_{a_n}^{b_n} \ldots \int_{a_1}^{b_1} p(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

The univariate or marginal distribution of X_j can be computed from the joint density by integrating out the variable not of interest.

$$\mathrm{P}(a_j \le X_j \le b_j) = \int_{-\infty}^{\infty} \ldots \int_{a_j}^{b_j} \ldots \int_{-\infty}^{\infty} p(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

The intuitive notion of *independence* of two random variables X_1, X_2 is formalized by requiring:

$$\mathrm{P}(a_1 \le X_1 \le b_1,\, a_2 \le X_2 \le b_2) = \mathrm{P}(a_1 \le X_1 \le b_1)\, \cdot\, \mathrm{P}(a_2 \le X_2 \le b_2),$$

i.e. the joint probability of two events depending on the random vector (X_1, X_2) can be factorized. It is sufficient to consider the univariate distributions of X_1 and X_2 exclusively. If the random vector (X_1, X_2) has a density $p(x_1, x_2)$, then X_1 and X_2 have densities $p_1(x)$ and $p_2(x)$ as well. In this case, independence of both random variables is equivalent to a joint density which can be factorized:

$$p(x_1, x_2) = p_1(x_1)p_2(x_2).$$

Dependence of two random variables X_1, X_2 can be very complicated. If X_1, X_2 are jointly normally distributed, their dependency structure can be rather easily quantified by their covariance:

$$\mathrm{Cov}(X_1, X_2) = \mathsf{E}[(X_1 - \mathsf{E}[X_1])(X_2 - \mathsf{E}[X_2])],$$

as well as by their correlation:

$$\mathrm{Corr}(X_1, X_2) = \frac{\mathrm{Cov}(X_1, X_2)}{\sigma(X_1) \cdot \sigma(X_2)}.$$

The correlation has the advantage of taking values between -1 and +1, which is scale invariant. For jointly normally distributed random variables, independence is equivalent to zero correlation, while complete dependence is equivalent to either a correlation of +1 (X_1 is large when X_2 is large) or a correlation of -1 (X_1 is large when X_2 is small).

In general, it holds for *independent* random variables $X_1, \ldots, X_n$

$$\mathrm{Cov}(X_i, X_j) = 0 \qquad \text{for } i \neq j.$$

This implies a useful computation rule:

$$\mathrm{Var}\Big(\sum_{j=1}^{n} X_j\Big) = \sum_{j=1}^{n} \mathrm{Var}(X_j).$$

If $X_1, \ldots, X_n$ are independent and have all the same distribution:

$$\mathrm{P}(a \le X_i \le b) = \mathrm{P}(a \le X_j \le b) \qquad \text{for all } i, j,$$

we call them *independently and identically distributed (i.i.d.)*.

3.5 Conditional Probabilities and Expectations

The *conditional probability* that a random variable Y takes values between a and b conditioned on the event that a random variable X takes values between x and $x + \Delta_x$, is defined as

$$\mathrm{P}(a \le Y \le b | x \le X \le x + \Delta_x) = \frac{\mathrm{P}(a \le Y \le b\,,\, x \le X \le x + \Delta_x)}{\mathrm{P}(x \le X \le x + \Delta_x)}\,, \quad (3.4)$$

provided the denominator is different from zero. The conditional probability of events of the kind $a \le Y \le b$ reflects our opinion of which values are more plausible than others, given that another random variable X has taken a certain value. If Y is independent of X, the probabilities of Y are not influenced by a priori knowledge about X. It holds:

$$\mathrm{P}(a \le Y \le b | x \le X \le x + \Delta x) = \mathrm{P}(a \le Y \le b)\,.$$

As Δx goes to 0 in equation (3.4), the left side of equation (3.4) converges heuristically to $\mathrm{P}(a \le Y \le b | X = x)$. In the case of a continuous random variable X having a density p_X, the left side of equation (3.4) is not defined since $\mathrm{P}(X = x) = 0$ for all x. But, it is possible to give a sound mathematical definition of the conditional distribution of Y given $X = x$. If the random variables Y and X have a joint distribution $p(x, y)$, then the conditional distribution has the density

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)} \quad \text{for} \quad p_X(x) \ne 0$$

and $p_{Y|X}(y|x) = 0$ otherwise. Consequently, it holds

$$\mathrm{P}(a \le Y \le b | X = x) = \int_a^b p_{Y|X}(y|x) dy.$$

The expectation with respect to the conditional distribution can be computed by

$$\mathsf{E}(Y|X = x) = \int_{-\infty}^{\infty} y\, p_{Y|X}(y|x) dy \stackrel{\text{def}}{=} \eta(x).$$

The function $\eta(x) = \mathsf{E}(Y|X = x)$ is called the *conditional expectation of* Y *given* $X = x$. Intuitively, it is the expectation of the random variable Y knowing that X has taken the value x.

Considering $\eta(x)$ as a function of the random variable X the conditional expectation of Y given X is obtained:

$$\mathsf{E}(Y|X) = \eta(X).$$

$\mathsf{E}(Y|X)$ is a random variable, which can be regarded as a function having the same expectation as Y. The conditional expectation has some useful properties, which we summarize in the following theorem.

Theorem 3.1 *Let X, Y, Z be real valued continuous random variables having a joint density.*

a) If X, Y are independent, then $\mathsf{E}(Y|X = x) = \mathsf{E}(Y)$

b) If $Y = g(X)$ is a function of X, then

$$\mathsf{E}[Y|X = x] = \mathsf{E}[g(X)|X = x] = g(x).$$

In general, it holds for random variables of the kind $Y = Zg(X)$:

$$\mathsf{E}[Y|X = x] = \mathsf{E}[Zg(X)|X = x] = g(x)\,\mathsf{E}[Z|X = x].$$

c) The conditional expectation is linear, i.e. for any real numbers a, b it holds:

$$\mathsf{E}(aY + bZ|X = x) = a\,\mathsf{E}(Y|X = x) + b\,\mathsf{E}(Z|X = x).$$

d) The law of iterated expectations: $\mathsf{E}[\mathsf{E}(Y|X)] = \mathsf{E}(Y)$.

The concept of the conditional expectation can be generalized analogously for multivariate random vectors Y and X. Let $S_t, t = 0, 1, 2, ...$ be a sequence of chronologically ordered random variables, for instance as a model of daily stock prices, let $Y = S_{t+1}$ and $X = (S_t, ..., S_{t-p+1})^\top$, then the conditional expectation

$$\mathsf{E}(Y|X = x) = \mathsf{E}(S_{t+1}|S_t = x_1, ..., S_{t-p+1} = x_p)$$

represents the expected stock price of the following day $t + 1$ given the stock prices $x = (x_1, ..., x_p)^\top$ of the previous p days. Since the information available at time t (relevant for the future evolution of the stock price) can consist of more than only a few past stock prices, we make frequent use of the notation $\mathsf{E}(Y|\mathcal{F}_t)$ for the expectation of Y given the information available up to time t. For all t, $\mathcal{F}_t$ denotes a family of events (having the structure of a $\sigma-$algebra, i.e. certain combinations of events of $\mathcal{F}_t$ are again elements of $\mathcal{F}_t$) representing the information available up to time t. $\mathcal{F}_t$ consists of events of which it is known whether they occur up to time t or not. Since more information unveils as time evolves, we must have $\mathcal{F}_s \subset \mathcal{F}_t$ for $s < t$, see Definition 5.1. Leaving out the exact definition of $\mathsf{E}(Y|\mathcal{F}_t)$ we confine to emphasize that the computation rules given in Theorem 3.1, appropriately reformulated, can be applied to the general conditional expectation.

3.6 Recommended Literature

Ross (1994), Pitman (1997), Krengel (2000) and Krengel (1995), among others, give an introduction to probability theory. An introduction to martingale theory which is imperative for the understanding of advanced mathematical finance is given by Williams (1991).

4 Stochastic Processes in Discrete Time

A *stochastic process* or random process consists of chronologically ordered random variables $\{X_t;\ t \geq 0\}$. For simplicity we assume that the process starts at time $t = 0$ in $X_0 = 0$. In this chapter, we consider exclusively processes in *discrete time*, i.e. processes which are observed at equally spaced points of time $t = 0, 1, 2, \dots$. Typical examples are daily, monthly or yearly observed economic data as stock prices, rates of unemployment or sales amount.

4.1 Binomial Processes

One of the simplest stochastic processes is an *ordinary random walk*, a process whose increments $Z_t = X_t - X_{t-1}$ from time $t-1$ to time t take exclusively the values +1 or -1. Additionally, we assume the increments to be i.i.d. and independent of the starting value X_0. Hence, the ordinary random walk can be written as:

$$X_t = X_0 + \sum_{k=1}^{t} Z_k \qquad , \ t = 1, 2, \dots \tag{4.1}$$

$X_0, Z_1, Z_2, \dots$ independent and

$$\mathrm{P}(Z_k = 1) = p \quad , \quad \mathrm{P}(Z_k = -1) = 1 - p \qquad \text{for all } \ k\,.$$

Letting the process go up by u and go down by d, instead, we obtain a more general class of *binomial processes*:

$$\mathrm{P}(Z_k = u) = p\,, \quad \mathrm{P}(Z_k = -d) = 1 - p \quad \text{für alle } \ k\,,$$

where u and d are constant (u=up, d=down).

Linear interpolation of the points (t, X_t) reflects the time evolution of the process and is called a *path* of an ordinary random walk. Starting in $X_0 = a$, the process moves on the grid of points (t, b_t), $t = 0, 1, 2, \ldots$, $b_t = a - t$, $a - t + 1, \ldots, a + t$. Up to time t, X_t can grow at most up to $a + t$ (if $Z_1 = \ldots = Z_t = 1$) or can fall at least to $a - t$ (if $Z_1 = \ldots = Z_t = -1$). Three paths of an ordinary random walk are shown in Figure 4.1 ($p = 0.5$), 4.2 ($p = 0.4$) and Figure 4.3 ($p = 0.6$).

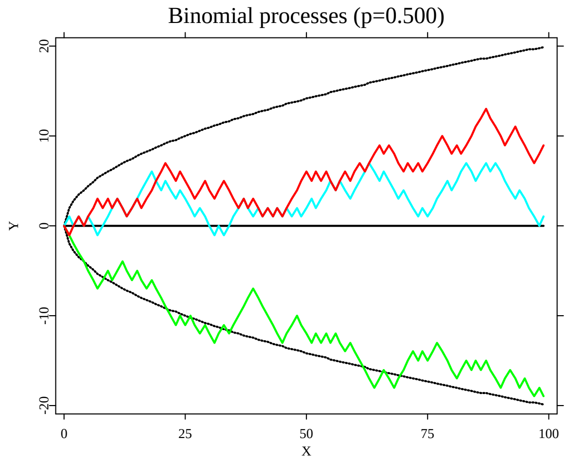


Figure 4.1: Three paths of a symmetric ordinary random walk. (2σ)–intervals around the drift (which is zero) are given as well. SFEBinomp.xpl

For generalized binomial processes the grid of possible paths is more complicated. The values which the process X_t starting in a can possibly take up to time t are given by

$$b_t = a + n \cdot u - m \cdot d\,, \text{ where } n, m \geq 0\,, \quad n + m = t\,.$$

If, from time 0 to time t, the process goes up n times and goes down m times then $X_t = a + n \cdot u - m \cdot d$. That is, n of t increments $Z_1, \ldots, Z_t$ take the value u, and m increments take the value $-d$. The grid of possible paths is also called a binomial tree.

The mean of the *symmetric ordinary random walk* ($p = \frac{1}{2}$) starting in 0 ($X_0 = 0$) is for all times t equal to 0 :

$$\mathsf{E}[X_t] = 0 \qquad \text{for all } \; t\;.$$

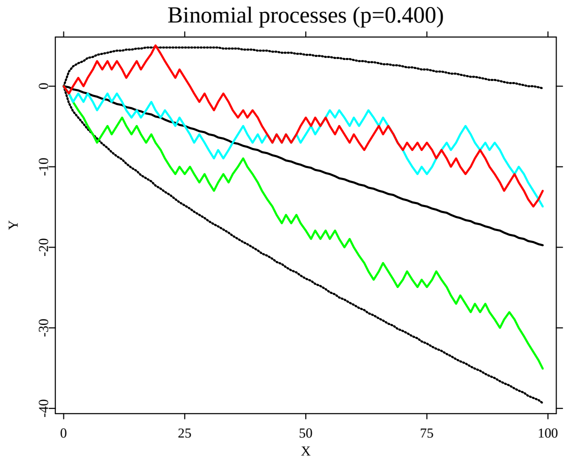


Figure 4.2: Three paths of an ordinary random walk with $p = 0.4$. (2σ)–intervals around the drift (which is the line with negative slope) are given as well. SFEBinomp.xpl

Otherwise, the *random walk* has a *trend* or *drift*, for $(p > \frac{1}{2})$ it has a positive drift and for $(p < \frac{1}{2})$ it has a negative drift. The process grows or falls in average:

$$\mathsf{E}[X_t] = t \cdot (2p - 1)\,,$$

since it holds for all increments $\mathsf{E}[Z_k] = 2p - 1$. Hence, the trend is linear in time. It is the upward sloping line in Figure 4.3 $(p = 0.6)$ and the downward sloping line in Figure 4.2 $(p = 0.6)$.

For the generalized *binomial process* with general starting value X_0 it holds analogously $\mathsf{E}[Z_k] = (u + d)p - d$ and thus:

$$\mathsf{E}[X_t] = \mathsf{E}[X_0] + t \cdot \{(u + d)p - d\}\,.$$

As time evolves the set of values X_t grows, and its variability increases. Since the summands in (4.1) are independent and $\mathrm{Var}(Z_k) = \mathrm{Var}(Z_1)$ for all k, the variance of X_t is given by (refer to Section 3.4):

$$\mathrm{Var}(X_t) = \mathrm{Var}(X_0) + t \cdot \mathrm{Var}(Z_1)\,.$$

Hence, the variance of X_t grows linearly with time. So does the standard deviation. For the random walks depicted in Figure 4.1 $(p = 0.5)$, Figure

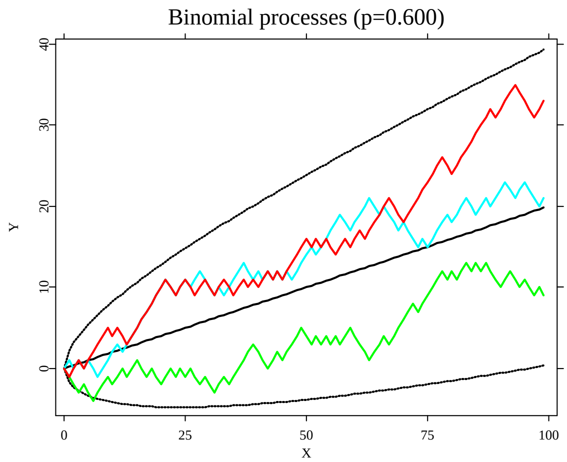


Figure 4.3: Three paths of an ordinary random walk with $p = 0.6$. (2σ)–intervals around the drift (which is the line with positive slope) are given as well. SFEBinomp.xpl

4.2 ($p = 0.4$) and Figure 4.3 ($p = 0.6$) the intervals $[\mathsf{E}[X_t] - 2\sigma(X_t); \mathsf{E}[X_t] + 2\sigma(X_t)]$ are shown as well. For large t , these intervals should contain 95% of the realizations of processes.

The variance of the increments can be easily computed. We use the following result which holds for the binomial distribution. Define

$$Y_k = \frac{Z_k + d}{u + d} = \begin{cases} 1 & \text{if } Z_k = u \\ 0 & \text{if } Z_k = -d \end{cases}$$

or

$$Z_k = (u + d)\, Y_k - d \tag{4.2}$$

we obtain the following representation of the binomial process

$$X_t = X_0 + (u + d)\, B_t - td \tag{4.3}$$

where

$$B_t = \sum_{k=1}^{t} Y_k \tag{4.4}$$

is a $B(t, p)$ distributed random variable.

Given the distribution of X_0, the distribution of X_t is specified for all t. It can be derived by means of a simple transformation of the binomial distribution $B(t,p)$. From equations (4.2) to (4.4) we obtain for $X_0 = 0$:

$$\mathrm{Var}(X_t) = t(u+d)^2\, p(1-p)$$

and for large t the distribution of X_t can be approximated by:

$$\mathcal{L}(X_t) \approx N(t\{(u+d)p - d\},\, t(u+d)^2\, p(1-p)).$$

For $p = \frac{1}{2}, u = d = \Delta x$, the following approximation holds for $\mathcal{L}(X_t)$:

$$N(0, t \cdot (\Delta x)^2).$$

Figure 4.4 shows the fit for $t = 100$.

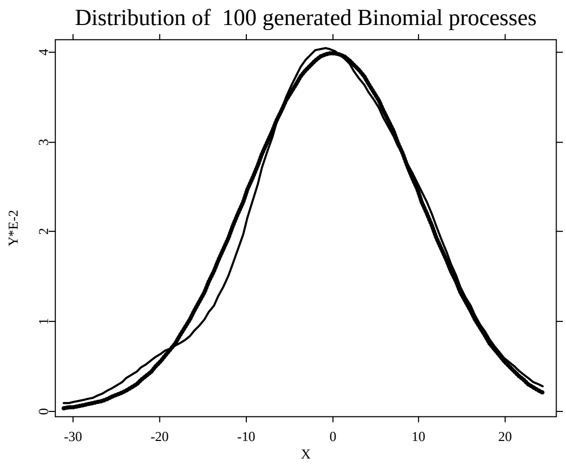


Figure 4.4: The distribution of 100 paths of an ordinary symmetric random walk of length 100 and a kernel density estimation of 100 normally distributed random variables. SFEbinomv.xpl

4.2 Trinomial Processes

In contrast to binomial processes, a *trinomial process* allows a quantity to stay constant within a given period of time. In the latter case, the increments are described by:

$$\mathrm{P}(Z_k = u) = p\,,\ \mathrm{P}(Z_k = -d) = q\,,\ \mathrm{P}(Z_k = 0) = r = 1 - p - q\ ,$$

and the process X_t is again given by:

$$X_t = X_0 + \sum_{k=1}^{t} Z_k$$

where $X_0, Z_1, Z_2, \ldots$ are mutually independent. To solve the Black–Scholes equation, some algorithms use trinomial schemes with time and state dependent probabilities p, q and r. Figure 4.5 shows five simulated paths of a trinomial process with $u = d = 1$ and $p = q = 0.25$.

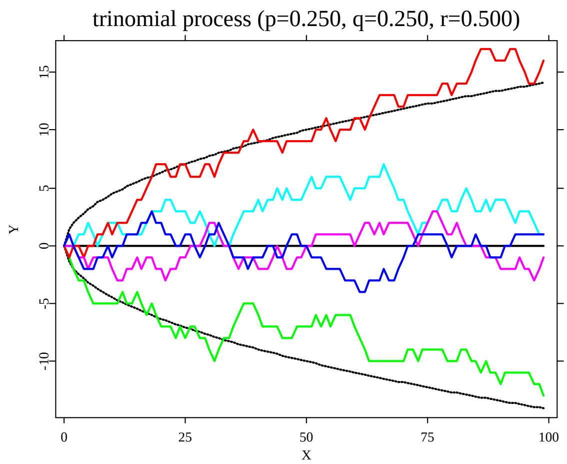


Figure 4.5: Five paths of a trinomial process with $p = q = 0.25$. (2σ)–intervals around the trend (which is zero) are given as well. SFETrinomp.xpl

The exact distribution of X_t cannot be derived from the binomial distribution but for the trinomial process a similar relations hold:

$$\begin{aligned}
\mathsf{E}[X_t] &= \mathsf{E}[X_0] + t \cdot \mathsf{E}[Z_1] = \mathsf{E}[X_0] + t \cdot (pu - qd) \\
\mathrm{Var}(X_t) &= \mathrm{Var}(X_0) + t \cdot \mathrm{Var}(Z_1), \text{ where} \\
\mathrm{Var}(Z_1) &= p(1-p)u^2 + q(1-q)d^2 + 2pq\ ud\ .
\end{aligned}$$

For large t, X_t is approximately $\mathrm{N}(\mathsf{E}[X_t], \mathrm{Var}(X_t))$–distributed.

4.3 General Random Walks

Binomial and trinomial processes are simple examples for *general random walks*, i.e. stochastic processes $\{X_t;\, t \geq 0\}$ satisfying:

$$X_t = X_0 + \sum_{k=1}^{t} Z_k\,, \qquad t = 1, 2, \ldots$$

where X_0 is independent of $Z_1, Z_2, \ldots$ which are i.i.d. The increments have a distribution of a real valued random variable. Z_k can take a finite or countably infinite number of values; but it is also possible for Z_k to take values out of a continuous set.

As an example, consider a *Gaussian random walk* with $X_0 = 0$, where the finitely many $X_1, \ldots, X_t$ are jointly normally distributed. Such a random walk can be constructed by assuming identically, independently and normally distributed increments. By the properties of the normal distribution, it follows that X_t is $N(\mu t,\, \sigma^2 t)$–distributed for each t. If $X_0 = 0$ and $\mathrm{Var}(Z_1)$ is finite, it holds approximately for all random walks for t large enough:

$$\mathcal{L}(X_t) \approx N(t \cdot \mathsf{E}[Z_1],\; t \cdot \mathrm{Var}(Z_1)).$$

This result follows directly from the central limit theorem for i.i.d. random variables.

Random walks are processes with *independent increments*. That means, the increment Z_{t+1} of the process from time t to time $t+1$ is independent of the past values $X_0, \ldots, X_t$ up to time t. In general, it holds for any $s > 0$ that the increment of the process from time t to time $t+s$

$$X_{t+s} - X_t = Z_{t+1} + \ldots + Z_{t+s}$$

is independent of $X_0, \ldots, X_t$. It follows that the best prediction, in terms of mean squared error, for X_{t+1} given $X_0, \ldots, X_t$ is just $X_t + \mathsf{E}[Z_{t+1}]$. As long as the price of only one stock is considered, this prediction rule works quite well. Already hundred years ago, Bachelier postulated (assuming $\mathsf{E}[Z_k] = 0$ for all k): "The best prediction for the stock price of tomorrow is the price of today."

Processes with independent increments are also *Markov–processes*. In other words, the future evolution of the process in time t depends exclusively on X_t, and the value of X_t is independent of the past values $X_0, \ldots, X_{t-1}$. If the increments Z_k and the starting value X_0, and hence all X_t, can take a finite or countably infinite number of values, then the *Markov–property* is formally expressed by:

$$\mathrm{P}(a_{t+1} < X_{t+1} < b_{t+1} | X_t = c,\, a_{t-1} < X_{t-1} < b_{t-1}, \ldots,\, a_0 < X_0 < b_0)$$
$$= \mathrm{P}(a_{t+1} < X_{t+1} < b_{t+1} | X_t = c)\,.$$

If $X_t = c$ is known, additional information about $X_0, \ldots, X_{t-1}$ does not influence the opinion about the range in which X_t will probably fall.

4.4 Geometric Random Walks

The essential idea underlying the random walk for real processes is the assumption of mutually independent increments of the order of magnitude for each point of time. However, economic time series in particular do not satisfy the latter assumption. Seasonal fluctuations of monthly sales figures for example are in *absolute terms* significantly greater if the yearly average sales figure is high. By contrast, the relative or percentage changes are stable over time and do not depend on the current level of X_t. Analogously to the random walk with i.i.d. absolute increments $Z_t = X_t - X_{t-1}$, a *geometric random walk* $\{X_t;\, t \geq 0\}$ is assumed to have i.i.d. relative increments

$$R_t = \frac{X_t}{X_{t-1}}\,, \quad t = 1, 2, \ldots\,.$$

For example, a geometric *binomial random walk* is given by

$$X_t = R_t \cdot X_{t-1} = X_0 \cdot \Pi_{k=1}^{t} R_k \tag{4.5}$$

where $X_0, R_1, R_2, \ldots$ are mutually independent and for $u > 1$, $d < 1$:

$$\mathrm{P}(R_k = u) = p\ ,\ \mathrm{P}(R_k = d) = 1 - p\ .$$

Given the independence assumption and $\mathsf{E}[R_k] = (u-d)p + d$ it follows from equation (4.5) that $\mathsf{E}[X_t]$ increases or decreases exponentially as the case may be $\mathsf{E}[R_k] > 1$ or $\mathsf{E}[R_k] < 1$:

$$\mathsf{E}[X_t] = \mathsf{E}[X_0] \cdot (\mathsf{E}[R_1])^t = \mathsf{E}[X_0] \cdot \{(u-d)p + d\}^t.$$

If $\mathsf{E}[R_k] = 1$ the process is on average stable, which is the case for

$$p = \frac{1-d}{u-d}\,.$$

For a recombining process, i.e. $d = \frac{1}{u}$, this relationship simplifies to

$$p = \frac{1}{u+1}\,.$$

Taking logarithms in equation (4.5) yields:

$$\ln X_t = \ln X_0 + \sum_{k=1}^{t} \ln R_k \ .$$

The process $\tilde{X}_t = \ln X_t$ is itself an ordinary binomial process with starting value $\ln X_0$ and increments $Z_k = \ln R_k$ for which hold:

$$\mathrm{P}(Z_k = \ln u) = p, \quad \mathrm{P}(Z_k = \ln d) = 1 - p \ .$$

For t large, $\tilde{X}_t$ is approximately normally distributed, i.e. $X_t = \exp(\tilde{X}_t)$ is approximately lognormally distributed.

4.5 Binomial Models with State Dependent Increments

Binomial processes and more general random walks model the stock price at best locally. They proceed from the assumption that the distribution of the increments $Z_t = X_t - X_{t-1}$ are the same for each value X_t, regardless of whether the stock price is substantially greater or smaller than X_0. Absolute increments $X_t - X_{t-1} = (R_t - 1)\, X_{t-1}$ of a geometric random walk depend on the level of X_{t-1}. Thus, geometric random walks are processes which do not have independent absolute increments. Unfortunately, modelling the stock price dynamics globally the latter processes are too simple to explain the impact of the current price level on the future stock price evolution. A class of processes which take this effect into account are binomial processes with state dependent (and possibly time dependent) increments:

$$X_t = X_{t-1} + Z_t \ , \quad t = 1, 2, \ldots \tag{4.6}$$

$$\mathrm{P}(Z_t = u) = p(X_{t-1}, t)\, , \ \mathrm{P}(Z_t = -d) = 1 - p(X_{t-1}, t)\, .$$

Since the distribution of Z_t depends on the state X_{t-1} and possibly on time t, increments are neither independent nor identically distributed. The deterministic functions $p(x, t)$ associate a probability to each possible value of the process at time t and to each t. Stochastic processes $\{X_t;\ t \geq 0\}$ which are constructed as in (4.6) are still *markovian* but without having independent increments.

Accordingly, geometric binomial processes with state dependent relative increments can be defined (for $u > 1,\ d < 1$):

$$X_t = R_t \cdot X_{t-1} \tag{4.7}$$

$$\mathrm{P}(R_t = u) = p(X_{t-1}, t)\,, \;\; \mathrm{P}(R_t = d) = 1 - p(X_{t-1}, t)\,.$$

Processes as defined in (4.6) and (4.7) are mainly of theoretic interest, since without further assumptions it is rather difficult to estimate the probabilities $p(x, t)$ from observed stock prices. But generalized binomial models (as well as the trinomial models) can be used to solve differential equations numerically, as the Black–Scholes equation for American options for example.

5 Stochastic Integrals and Differential Equations

This chapter provides the tools which are used in option pricing. The field of stochastic processes in continuous time which are defined as solutions of stochastic differential equations plays an important role. To illustrate these notions we use repeatedly approximations by stochastic processes in discrete time and refer to the results from Chapter 4.

A stochastic process in continuous time $\{X_t;\ t \ge 0\}$ consists of chronologically ordered random variables, but here the variable t is continuous, i.e. t is a positive real number.

Stock prices are actually processes in discrete time. But to derive the Black-Scholes equation they are approximated by continuous time processes which are easier to handle analytically. However the simulation on a computer of such processes or the numerical computation of say American options, is carried out by means of discrete time approximations. We therefore switch the time scale between discrete and continuous depending on what is more convenient for the actual computation.

5.1 Wiener Process

We begin with a simple symmetric random walk $\{X_n;\ n \ge 0\}$ starting in 0 ($X_0 = 0$). The increments $Z_n = X_n - X_{n-1}$ are i.i.d. with :

$$\mathrm{P}(Z_n = 1) = \mathrm{P}(Z_n = -1) = \frac{1}{2}\,.$$

By shortening the period of time of two successive observations we accelerate the process. Simultaneously, the increments of the process become smaller during the shorter period of time. More precisely, we consider a stochastic process $\{X_t^\Delta;\ t \ge 0\}$ in continuous time which increases or decreases in a time step Δt with probability $\frac{1}{2}$ by Δx. Between these jumps the process is constant (alternatively we could interpolate linearly). At time $t = n \cdot \Delta t$ the process is:

$$X_t^\Delta = \sum_{k=1}^{n} Z_k \cdot \Delta x = X_n \cdot \Delta x$$

where the increments $Z_1 \Delta x, Z_2 \Delta x, \ldots$ are mutually independent and take the values Δx or $-\Delta x$ with probability $\frac{1}{2}$ respectively. From Section 4.1 we know:

$$\mathsf{E}[X_t^\Delta] = 0 \ , \qquad \mathrm{Var}(X_t^\Delta) = (\Delta x)^2 \cdot \mathrm{Var}(X_n) = (\Delta x)^2 \cdot n = t \cdot \frac{(\Delta x)^2}{\Delta t} \ .$$

Now, we let Δt and Δx become smaller. For the process in the limit to exist in a reasonable sense, $\mathrm{Var}(X_t^\Delta)$ must be finite. On the other hand, $\mathrm{Var}(X_t^\Delta)$ should not converge to 0, since the process would then not be random any more. Hence, we must choose:

$$\Delta t \to 0, \ \Delta x = c \cdot \sqrt{\Delta t}\,, \text{ such that } \ \mathrm{Var}(X_t^\Delta) \to c^2 t\,.$$

If Δt is small, then $n = t/\Delta t$ is large. Thus, the random variable X_n of the ordinary symmetric random walk is approximately $\mathrm{N}(0, n)$ distributed, and therefore for all t (not only for t such that $t = n \ \Delta t$):

$$\mathcal{L}(X_t^\Delta) \approx \mathrm{N}(0, \ n(\Delta x)^2) \approx \mathrm{N}(0, \ c^2 t) \ .$$

Thus the limiting process $\{X_t; \ t \geq 0\}$ which we obtain from $\{X_t^\Delta; \ t \geq 0\}$ for $\Delta t \to 0, \ \Delta x = c \ \sqrt{\Delta t}$ has the following properties:

(i) X_t is $\mathrm{N}(0, c^2 t)$ distributed for all $t \geq 0$.

(ii) $\{X_t; \ t \geq 0\}$ has *independent increments*, i.e. for $0 \leq s < t$, $X_t - X_s$ is independent of X_s (since the random walk $\{X_n; \ n \geq 0\}$ defining $\{X_t^\Delta; \ t \geq 0\}$ has independent increments).

(iii) For $0 \leq s < t$ the increment $(X_t - X_s)$ is $\mathrm{N}(0, \ c^2 \cdot (t - s))$ distributed, i.e. its distribution depends exclusively on the length $t - s$ of the time interval in which the increment is observed (this follows from (i) and (ii) and the properties of the normal distribution).

A stochastic process $\{X_t; \ t \geq 0\}$ in continuous time satisfying (i)–(iii) is called *Wiener process* or *Brownian motion* starting in 0 ($X_0 = 0$). The standard Wiener process resulting from $c = 1$ will be denoted by $\{W_t; \ t \geq 0\}$. For this process it holds for all $0 \leq s < t$

$$\mathsf{E}[W_t] = 0, \ \mathrm{Var}(W_t) = t$$

$$\begin{aligned}\mathrm{Cov}(W_t, W_s) &= \mathrm{Cov}(W_t - W_s + W_s, W_s)\\ &= \mathrm{Cov}(W_t - W_s, W_s) + \mathrm{Cov}(W_s, W_s)\\ &= 0 + \mathrm{Var}(W_s) = s\end{aligned}$$

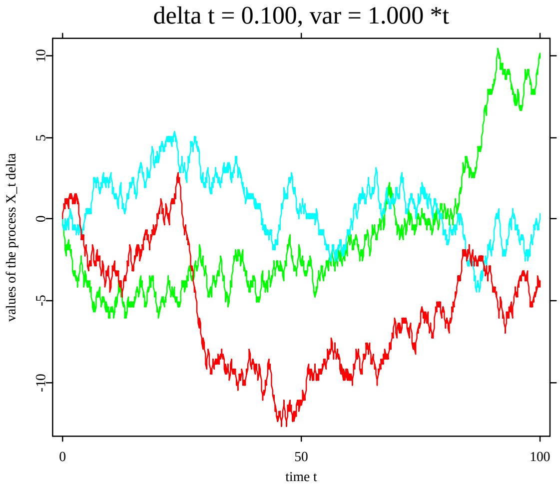


Figure 5.1: Typical paths of a Wiener process. SFEWienerProcess.xpl

As for every stochastic process in continuous time, we can consider a path or realization of the Wiener process as a *randomly chosen function* of time. With some major mathematical instruments it is possible to show that the paths of a Wiener process are continuous with probability 1:

$$\mathrm{P}(W_t \text{ is continuous as a function of } t) = 1\,.$$

That is to say, the Wiener process has no jumps. But W_t fluctuates heavily: the paths are continuous but highly erratic. In fact, it is possible to show that the paths are not differentiable with probability 1.

Being a process with independent increments the Wiener process is *markovian*. For $0 \le s < t$ it holds $W_t = W_s + (W_t - W_s)$, i.e. W_t depends only on W_s and on the increment from time s to time t:

$$\begin{aligned}&\mathrm{P}(a < W_t < b\,|W_s = x\ ,\ \text{information about } \ W_{t'},\ 0 \le t' < s)\\ =\ &\mathrm{P}(a < W_t < b\,|W_s = x\ )\end{aligned}$$

Using properties (i)–(iii), the distribution of W_t given the outcome $W_s = x$ can be formulated explicitly. Since the increment $(W_t - W_s)$ is $\mathrm{N}(0, t-s)$ distributed, W_t is $\mathrm{N}(x, t-s)$ distributed given $W_s = x$:

$$\mathrm{P}(a < W_t < b \,|W_s = x) = \int_a^b \frac{1}{\sqrt{t-s}}\varphi\Big(\frac{y-x}{\sqrt{t-s}}\Big)dy.$$

Proceeding from the assumption of a random walk $\{X_n;\, n \geq 0\}$ with drift $\mathsf{E}[X_n] = n(2p-1)$ instead of a symmetric random walk results in a process X_t^Δ which is no more zero on average, but

$$\begin{aligned}\mathsf{E}[X_t^\Delta] &= n\cdot(2p-1)\cdot\Delta x = (2p-1)\cdot t\,\frac{\Delta x}{\Delta t}\\ \mathrm{Var}(X_t^\Delta) &= n\; 4p(1-p)\cdot(\Delta x)^2 = 4p(1-p)\cdot t\cdot\frac{(\Delta x)^2}{\Delta t}.\end{aligned}$$

For $\Delta t \to 0$, $\Delta x = \sqrt{\Delta t}$, $p = \frac{1}{2}(1+\mu\sqrt{\Delta t})$ we obtain for all t:

$$\mathsf{E}[X_t^\Delta] \to \mu t\,,\ \mathrm{Var}(X_t^\Delta) \to t.$$

The limiting process is a Wiener process $\{X_t;\ t \geq 0\}$ with *drift* or *trend* μt. It results from the standard Wiener process:

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

Hence, it behaves in the same way as the standard Wiener process but it fluctuates on average around μ instead of 0. If $(\mu > 0)$ the process is increasing linearly on average, and if $(\mu < 0)$ it is decreasing linearly on average.

5.2 Stochastic Integration

In order to introduce a stochastic process as a solution of a stochastic differential equation as we introduce the concept of the Itô–integral: a stochastic integral with respect to a Wiener process. Formally the construction of the Itô–integral is similar to the one of the Stieltjes–integral. But instead of integrating with respect to a deterministic function (Stieltjes–integral), the Itô–integral integrates with respect to a random function, more precisely, the path of a Wiener process. Since the integrant itself can be random, i.e. it can be a path of a stochastic process, one has to analyze the mutual dependencies of the integrant and the Wiener process.

Let $\{Y_t;\ t \geq 0\}$ be the process to integrate and let $\{W_t;\ t \geq 0\}$ be a standard Wiener process. The definition of a stochastic integral assumes that

$\{Y_t;\ t \geq 0\}$ is not anticipating. Intuitively, it means that the process up to time s does not contain any information about future increments $W_t - W_s$, $t > s$, of the Wiener process. In particular, Y_s is independent of $W_t - W_s$.

An integral of a function is usually defined as the limit of the sum of the suitably weighted function. Similarly, the *Itô integral* with respect to a Wiener process is defined as the limit of the sum of the (randomly) weighted (random) function $\{Y_t;\ t \geq 0\}$:

$$I_n = \sum_{k=1}^{n} Y_{(k-1)\Delta t} \cdot (W_{k\Delta t} - W_{(k-1)\Delta t}) , \quad \Delta t = \frac{t}{n} \tag{5.1}$$

$$\int_0^t Y_s \, dW_s = \lim_{n\to\infty} I_n ,$$

where the limit is to be understood as the limit of a random variable in terms of mean squared error, i.e. it holds

$$\mathsf{E}\{\big[\int_0^t Y_s \, dW_s - I_n\big]^2\} \to 0, \quad n \to \infty.$$

It is important to note, that each summand of I_n is a product of two independent random variables. More precisely, $Y_{(k-1)\Delta t}$, the process to integrate at the left border of the small interval $[(k-1)\Delta t,\, k\Delta t]$ is independent of the increment $W_{k\Delta t} - W_{(k-1)\Delta t}$ of the Wiener process in this interval.

It is not hard to be more precise on the non anticipating property of $\{Y_t;\ t \geq 0\}$.

Definition 5.1 (Information structure, non–anticipating)
For each $t \geq 0$, $\mathcal{F}_t$ denotes a family of events (having the structure of a σ–algebra, i.e. certain combinations of events contained in $\mathcal{F}_t$ are again in $\mathcal{F}_t$) which contain the available information up to time t. $\mathcal{F}_t$ consists of events from which is known up to time t whether they occurred or not. We assume:

$\mathcal{F}_s \subset \mathcal{F}_t$	*for $s < t$*	*(information grows as time evolves)*
$\{a < Y_t < b\} \in \mathcal{F}_t$		*(Y_t contains no information about events occurring after time t)*
$\{a < W_t < b\} \in \mathcal{F}_t$		
$W_t - W_s$ *independent of* $\mathcal{F}_s$	*for $s < t$*	*(the Wiener process is adapted to evolution of information)*

Then, we call $\mathcal{F}_t$ the information structure at time t and the process $\{Y_t;\ t \geq 0\}$ non–anticipating with respect to the information structure $\mathcal{F}_t;\ t \geq 0$.

The process $\{Y_t\}$ is called non–anticipating since due to the second assumption it does not anticipate any future information. The evolving information structure $\mathcal{F}_t$ and the random variables Y_t, W_t are adapted to each other.

The integral depends crucially on the point of the interval $[(k-1)\Delta t,\ k\Delta t]$ at which the random variable Y_s is evaluated in (5.1). Consider the example $Y_t = W_t,\ t \geq 0$, i.e. we integrate the Wiener process with respect to itself. As a gedankenexperiment we replace in (5.1) $(k-1)\Delta t$ by an arbitrary point $t(n,k)$ of the interval $[(k-1)\Delta t,\ k\Delta t]$. If we defined:

$$\int_0^t W_s\, dW_s = \lim_{n\to\infty} \sum_{k=1}^{n} W_{t(n,k)}\,(W_{k\Delta t} - W_{(k-1)\Delta t})$$

the expected values would converge as well. Hence by interchanging the sum with the covariance operator we get:

$$\begin{aligned} \mathsf{E} \quad & [\sum_{k=1}^{n} W_{t(n,k)}\,(W_{k\Delta t} - W_{(k-1)\Delta t})] = \sum_{k=1}^{n} \mathrm{Cov}(W_{t(n,k)},\, W_{k\Delta t} - W_{(k-1)\Delta t}) \\ = \quad & \sum_{k=1}^{n} \{t(n,k) - (k-1)\Delta t\} \to \mathsf{E}[\int_0^t W_s\, dW_s]\,. \end{aligned}$$

For $t(n,k) = (k-1)\Delta t$ – which is the case for the Itô–integral – we obtain 0, for $t(n,k) = k\Delta t$ we obtain $n \cdot \Delta t = t$, and for suitably chosen sequences $t(n,k)$ we could obtain for the expectation of the stochastic integral any value between 0 and t. In order to assign to $\int_0^t W_s\ dW_s$ a unique value, we have to agree on a certain sequence $t(n,k)$.

To illustrate how Itô–integrals are computed, and that other than the usual computation rules have to be applied, we show that:

$$\int_0^t W_s\ dW_s = \frac{1}{2}(W_t^2 - W_0^2)\ -\ \frac{t}{2} = \frac{1}{2}(W_t^2 - t) \tag{5.2}$$

Summing the differences $W_{k\Delta t}^2 - W_{(k-1)\Delta t}^2$, all terms but the first and the last cancel out and remembering that $n\Delta t = t$ we get

$$\begin{aligned}
\frac{1}{2}(W_t^2 - W_0^2) &= \frac{1}{2}\sum_{k=1}^{n}(W_{k\Delta t}^2 - W_{(k-1)\Delta t}^2) \\
&= \frac{1}{2}\sum_{k=1}^{n}(W_{k\Delta t} - W_{(k-1)\Delta t})(W_{k\Delta t} + W_{(k-1)\Delta t}) \\
&= \frac{1}{2}\sum_{k=1}^{n}(W_{k\Delta t} - W_{(k-1)\Delta t})^2 \\
&\quad + \sum_{k=1}^{n}(W_{k\Delta t} - W_{(k-1)\Delta t})\, W_{(k-1)\Delta t}\,.
\end{aligned}$$

While the second term converges to $\int_0^t W_s\ dW_s$, the first term is a sum of n independent identically distributed random variables and which is thus approximated due to the law of large numbers by its expected value

$$\frac{n}{2}\,\mathsf{E}[(W_{k\Delta t} - W_{k-1)\Delta t})^2] = \frac{n}{2}\Delta t = \frac{t}{2}\,.$$

For smooth functions f_s, for example continuously differentiable functions, it holds $\int_0^t f_s\ df_s = \frac{1}{2}(f_t^2 - f_0^2)$. However, the stochastic integral (5.2) contains the additional term $-\frac{t}{2}$ since the local increment of the Wiener process over an interval of length Δt is of the size of its standard deviation – that is $\sqrt{\Delta t}$. The increment of a smooth function f_s is proportional to Δt, and therefore considerably smaller than the increment of the Wiener process for $\Delta t \to 0$.

5.3 Stochastic Differential Equations

Since the Wiener process fluctuates around its expectation 0 it can be approximated by means of symmetric random walks. As for random walks we are interested in stochastic processes in continuous time which are growing on average, i.e. which have a *trend* or *drift*. Proceeding from a Wiener process with arbitrary σ (see Section 5.1) we obtain the generalized Wiener process $\{X_t;\ t \geq 0\}$ with *drift rate* μ and variance σ^2 :

$$X_t = \mu \cdot t + \sigma \cdot W_t \quad , \quad t \geq 0\,. \tag{5.3}$$

The general Wiener process X_t is at time t, $\mathrm{N}(\mu t,\, \sigma^2 t)$–distributed. For its increment in a small time interval Δt we obtain

$$X_{t+\Delta t} - X_t = \mu \cdot \Delta t + \sigma(W_{t+\Delta t} - W_t)\,.$$

For $\Delta t \to 0$ use the differential notation:

$$dX_t = \mu \cdot dt + \sigma \cdot dW_t \tag{5.4}$$

This is only a different expression for the relationship (5.3) which we can also write in integral form:

$$X_t = \int_0^t \mu\, ds + \int_0^t \sigma dW_s \tag{5.5}$$

Note, that from the definition of the stochastic integral it follows directly that $\int_0^t d\ W_s = W_t - W_0 = W_t$.

The differential notation (5.4) proceeds from the assumption that both the local drift rate given by μ and the local variance given by σ^2 are constant. A considerably larger class of stochastic processes which is more suited to model numerous economic and natural processes is obtained if μ and σ^2 in (5.4) are allowed to be time and state dependent. Such processes $\{X_t;\ t \geq 0\}$, which we call Itô–processes, are defined as solutions of *stochastic differential equations*:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \tag{5.6}$$

Intuitively, this means:

$$X_{t+\Delta t} - X_t = \mu(X_t, t)\Delta t + \sigma(X_t, t)(W_{t+\Delta t} - W_t),$$

i.e. the process' increment in a small interval of length Δt after time t is $\mu(X_t, t)\cdot\Delta t$ plus a random fluctuation which is $\mathrm{N}(0,\ \sigma^2(X_t, t)\cdot\Delta t)$ distributed. A precise definition of a solution of (5.6) is a stochastic process fulfilling the integral equation

$$X_t - X_0 = \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s \tag{5.7}$$

In this sense (5.6) is only an abbreviation of (5.7). For $0 \leq t' < t$, it follows immediately:

$$X_t = X_{t'} + \int_{t'}^t \mu(X_s, s)ds + \int_{t'}^t \sigma(X_s, s)dW_s\,.$$

Since the increment of the Wiener process between t' and t does not dependent on the events which occurred up to time t', it follows that an Itô–process is *Markovian.*

Discrete approximations of (5.6) and (5.7) which can be used to simulate Itô–processes are obtained by observing the process between 0 and t only at evenly spaced points in time $k\Delta t,\ k = 0, \ldots, n\ ,\ \ n\Delta t = t$.

With $X_k = X_{k\Delta t}$ and $Z_k = (W_{k\Delta t} - W_{(k-1)\Delta t})/\sqrt{\Delta t}$ we get

$$X_{k+1} - X_k = \mu(X_k, k) \cdot \Delta t + \sigma(X_k, k) \cdot Z_{k+1} \cdot \sqrt{\Delta t}$$

or rather with the abbreviations $\mu_k(X) = \mu(X, k)\Delta t$, $\sigma_k(X) = \sigma(X, k)\sqrt{\Delta t}$:

$$X_n - X_0 = \sum_{k=1}^{n} \mu_{k-1}(X_{k-1}) + \sum_{k=1}^{n} \sigma_{k-1}(X_{k-1}) \cdot Z_k$$

with identical independently distributed $\mathrm{N}(0, 1)$–random variables $Z_1, Z_2, \ldots$.

`SFESimCIR.xpl`, `SFMSimOU.xpl`

5.4 The Stock Price as a Stochastic Process

Stock prices are stochastic processes in *discrete time* which take only *discrete values* due to the limited measurement scale. Nevertheless, stochastic processes in *continuous time* are used as models since they are analytically easier to handle than discrete models, e.g. the binomial or trinomial process. However, the latter are more intuitive and prove to be very useful in simulations.

Two features of the general Wiener process $dX_t = \mu dt + \sigma\, dW_t$ make it an unsuitable model for stock prices. First, it allows for negative stock prices, and second the local variability is higher for high stock prices. Hence, stock prices S_t are modeled by means of the more general Itô–process:

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t \, .$$

This model does depend on the unknown functions $\mu(X, t)$ and $\sigma(X, t)$. A useful and simpler variant utilizing only two unknown real model parameters μ and σ can be justified by the following reflection: The percentage return on the invested capital should on average not depend on the stock price at which the investment is done, and of course, should not depend on the *currency unit* (EUR, USD, ...) in which the stock price is quoted. Furthermore, the average return should be proportional to the investment horizon, as it is the case for other investment instruments. Putting things together, we request:

$$\frac{\mathsf{E}[dS_t]}{S_t} = \frac{\mathsf{E}[S_{t+dt} - S_t]}{S_t} = \mu \cdot dt \, .$$

Since $\mathsf{E}[dW_t] = 0$ this condition is satisfied if

$$\mu(S_t, t) = \mu \cdot S_t \, ,$$

for given S_t. Additionally,

$$\sigma(S_t, t) = \sigma \cdot S_t$$

takes into consideration that the absolute size of the stock price fluctuation is proportional to the currency unit in which the stock price is quoted. In summary, we model the stock price S_t as a solution of the stochastic differential equation

$$dS_t = \mu \cdot S_t \, dt + \sigma \cdot S_t \cdot dW_t \, ,$$

where μ is the *expected return* on the stock, and σ the *volatility.* Such a process is called *geometric Brownian motion* because

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

By applying Itôs lemma, which we introduce in Section 5.5, it can be shown that for a suitable Wiener process $\{Y_t;\ t \geq 0\}$ it holds

$$S_t = e^{Y_t} \quad \text{bzw.} \quad Y_t = \ln S_t \, .$$

Since Y_t is normally distributed, S_t is lognormally distributed. As random walks can be used to approximate the general Wiener process, geometric random walks can be used to approximate geometric Brownian motion and thus this simple model for the stock price.

5.5 Itôs Lemma

A crucial tool in dealing with stochastic differential equations is Itôs lemma. If $\{X_t,\ t \geq 0\}$ is an Itô–process:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t) \, dW_t \, , \tag{5.8}$$

one is often interested in the dynamics of stochastic processes which are functions of X_t: $Y_t = g(X_t)$. Then $\{Y_t;\ t \geq 0\}$ can also be described by a solution of a stochastic differential equation from which interesting properties of Y_t can be derived as for example the average growth in time t.

For a heuristic derivation of the equation for $\{Y_t;\ t \geq 0\}$ we assume that g is differentiable as many times as necessary. From a Taylor expansion it follows:

$$\begin{aligned} Y_{t+dt} - Y_t &= g(X_{t+dt}) - g(X_t) \\ &= g(X_t + dX_t) - g(X_t) \\ &= \frac{dg}{dX}(X_t) \cdot dX_t + \frac{1}{2}\frac{d^2 g}{dX^2}(X_t) \cdot (dX_t)^2 + \ldots \end{aligned} \tag{5.9}$$

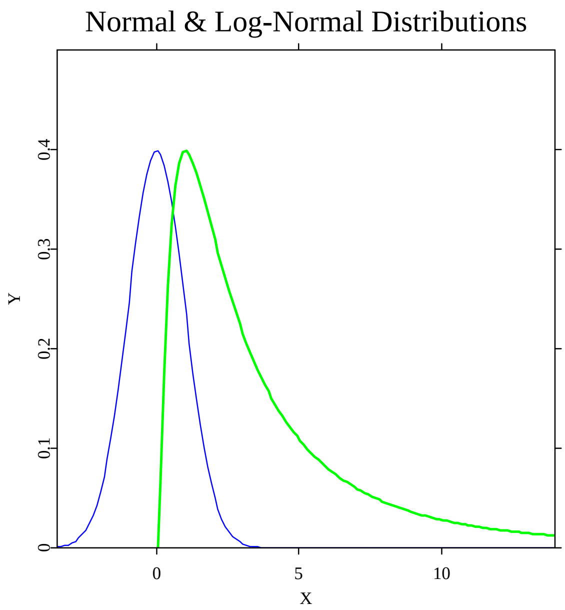


Figure 5.2: Density comparison of lognormally and normally distributed random variables. `SFELogNormal.xpl`

where the dots indicate the terms which can be neglected (for $dt \to 0$). Due to equation (5.8) the drift term $\mu(X_t, t)dt$ and the volatility term $\sigma(X_t, t)dW_t$ are the dominant terms since for $dt \to 0$ they are of size dt and $\sqrt{dt}$ respectively.

In doing this, we use the fact that $\mathsf{E}[(dW_t)^2] = dt$ and $dW_t = W_{t+dt} - W_t$ is of the size of its standard deviation, $\sqrt{dt}$. We neglect terms which are of smaller size than dt. Thus, we can express $(dX_t)^2$ by a simpler term:

$$\begin{aligned}(dX_t)^2 &= (\mu(X_t, t)dt + \sigma(X_t, t)\, dW_t)^2 \\ &= \mu^2(X_t, t)(dt)^2 + 2\mu(X_t, t)\ \sigma(X_t, t)\ dt\ dW_t + \sigma^2(X_t, t)(dW_t)^2\,.\end{aligned}$$

We see that the first and the second term are of size $(dt)^2$ and $dt \cdot \sqrt{dt}$ respectively. Therefore, both can be neglected. However, the third term is of size dt. More precisely, it can be shown that $dt \to 0$:

$$(dW_t)^2 = dt.$$

Thanks to this identity, calculus rules for stochastic integrals can be derived from the rules for deterministic functions (as Taylor expansions for example). Neglecting terms which are of smaller size than dt we obtain from (5.9) the following version of *Itôs lemma*:

Lemma 5.1 (Itôs Lemma)

$$\begin{aligned} dY_t &= dg(X_t) \\ &= \left(\frac{dg}{dX}(X_t) \cdot \mu(X_t, t) + \frac{1}{2}\frac{d^2 g}{dX^2}(X_t) \cdot \sigma^2(X_t, t)\right) dt \\ &\quad + \frac{dg}{dX}(X_t) \cdot \sigma(X_t, t)\, dW_t \end{aligned}$$

or - dropping the time index t and the argument X_t of the function g and its derivatives:

$$dg = \left(\frac{dg}{dX}\mu(X,t) + \frac{1}{2}\frac{d^2 g}{dX^2}\sigma^2(X,t)\right) dt + \frac{dg}{dX}\sigma(X,t) dW\,.$$

Example 5.1
Consider $Y_t = \ln S_t$ the logarithm of the geometric Brownian motion. For $g(X) = \ln X$ we obtain $\frac{dg}{dX} = \frac{1}{X}$, $\frac{d^2 g}{dX^2} = -\frac{1}{X^2}$. Applying Itôs lemma for the geometric Brownian motion with $\mu(X,t) = \mu X$, $\sigma(X,t) = \sigma X$ we get:

$$\begin{aligned} dY_t &= \left(\frac{1}{S_t}\mu S_t - \frac{1}{2}\frac{1}{S_t^2}\sigma^2 S_t^2\right)dt + \frac{1}{S_t} \cdot \sigma S_t\, dW_t \\ &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma\, dW_t \end{aligned}$$

The logarithm of the stock price *is a generalized Wiener process with drift rate $\mu^* = \mu - \frac{1}{2}\sigma^2$ and variance rate σ^2. Since Y_t is N($\mu^* t$, $\sigma^2 t$)–distributed S_t is itself* lognormally distributed *with parameters $\mu^* t$ and $\sigma^2 t$.*

A generalized version of Itôs lemma for functions $g(X,t)$ which are allowed to depend on time t is:

Lemma 5.2 (Itôs lemma for functions depending explicitly on time)

$$dg = \left(\frac{\partial g}{\partial X} \cdot \mu(X,t) + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}\sigma^2(X,t) + \frac{\partial g}{\partial t}\right) dt + \frac{\partial g}{\partial X}\sigma(X,t) dW \quad (5.10)$$

$Y_t = g(X_t, t)$ is again an Itô process, but this time the drift rate is augmented by an additional term $\frac{\partial g}{\partial t}(X_t, t)$.

5.6 Recommended Literature

This chapter briefly summarized results which belong to the main topics of stochastic analysis. Numerous textbooks of different levels introduce to the calculus of stochastic integrals and stochastic differential equations, see for example von Weizsäcker and Winkler (1990), Mikosch (1998) or Karatzas and Shreve (1999).

6 Black–Scholes Option Pricing Model

6.1 Black–Scholes Differential Equation

Simple generally accepted economic assumptions are insufficient to develop a rational option pricing theory. Assuming a perfect financial market in Section 2.1 lead to elementary arbitrage relations which options have to fulfill. While these relations can be used as a verification tool for sophisticated mathematical models, they do not provide an explicit option pricing function depending on parameters such as time and the stock price as well as the options underlying parameters K, T. To obtain such a pricing function the value of the underlying financial instrument (stock, currency, ...) has to be modelled. In general, the underlying instrument is assumed to follow a stochastic process either in discrete or in continuous time. While the latter are analytically easier to handle, the former, which we will consider as approximations of continuous time processes for the time being, are particularly useful for numerical computations. In the second part of this text, the discrete time version will be discussed as financial time series models.

A model for stock prices which is frequently used and is also the basis of the classical Black–Scholes approach, is the so–called geometric Brownian motion. In this model the stock price S_t is a solution of the stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t. \tag{6.1}$$

Equivalently, the process of stock price returns can be assumed to follow a standard Brownian motion, i.e.

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t. \tag{6.2}$$

The drift μ is the expected return on the stock in the time interval dt. The volatility σ is a measure of the return variability around its expectation μ.

Both parameters μ and σ are dependent on each other and are important factors of the investors' risk preferences involved in the investment decision: The higher the expected return μ, the higher, in general, the risk quantified by σ.

Modelling the underlying as geometric Brownian motion provides a useful approximation to stock prices accepted by practitioners for short and medium maturity. In real practice, numerous model departures are known: in some situations the volatility function $\sigma(x, t)$ of the general model (5.8) is different from the linear specification $\sigma \cdot x$ of geometric Brownian motion. The Black–Scholes' approach is still used to approximate option prices. The basic idea to derive option prices can be applied to more general stock price models.

Black–Scholes' approach relies on the idea introduced in Chapter 2, i.e. duplicating the portfolio consisting of the option by means of a second portfolio consisting exclusively of financial instruments whose values are known. The duplicating portfolio is chosen such that both portfolios have equal values at options maturity T. Then, it follows from the assumption of a perfect financial market, and in particular of no–arbitrage opportunities, that both portfolios must have equal values at any time prior to time T. The duplicating portfolio can be created in two equivalent ways which we illustrate in an example of a call option on a stock with price S_t:

1. Consider a portfolio consisting of one call of which the price is to be computed. The duplicating portfolio is composed of stocks and risk–less zero bonds of which the quantity adapts continuously to changes in the stock price. Without loss of generality, the zero bond's face value can be set equal to one since the number of zero bond's in the duplicating portfolio is free parameter. At time t the two portfolios consist of:

Portfolio A**:** One call option (long position) with delivery price K and maturity date T.

Portfolio B**:** $n_t = n(S_t, t)$ stocks and $m_t = m(S_t, t)$ zero bonds with face value $B_T = 1$ and maturity date T.

2. Consider a perfect hedge–portfolio, which consists of stocks and written calls (by means of short selling). Due to a dynamic hedge–strategy the portfolio bears no risk at any time, i.e. profits due to the calls are neutralized by losses due to the stocks. Correspondingly, the duplicating portfolio is also risk–less and consists exclusively of zero bonds. Again, the positions are adjusted continuously to changes in the stock price. At time t the two portfolios

are composed of:

Portfolio A**:** One stock and $n_t = n(S_t, t)$ (by means of short selling) written call options on the stock with delivery price K and maturity date T.

Portfolio B**:** $m_t = m(S_t, t)$ zero bonds with face value $B_T = 1$ and maturity dates T.

Let $T^* = T$ be the time when the call option expires worthless, and otherwise let T^* be the time at which the option is exercised. While for a European call option it holds $T^* = T$ at any time, an American option can be exercised prior to maturity. We will see that both in 1. the call value is equal to the value of the duplicating portfolio, and in 2. the hedge–portfolio's value equals the value of the risk–less zero bond portfolio at any time $t \leq T^*$, and thus the same partial differential equation for the call value results, which is called *Black–Scholes equation*.

The Black–Scholes approach can be applied to any financial instrument $\mathcal{U}$ contingent on an underlying with price S_t if the latter price follows a geometric Brownian motion, and if the derivatives price F_t is a function only of the price S_t and time: $F_t = F(S_t, t)$. Then, according to the theorem below, a portfolio duplicating the financial instrument exists, and the approach illustrated in 1. can be applied to price the instrument. Pricing an arbitrary derivative the duplicating portfolio must have not only the same value as the derivative at exercising time T^*, but also the same cash flow pattern, i.e. the duplicating portfolio has to generate equal amounts of withdrawal profits or contributing costs as the derivative. The existence of a perfect hedge–portfolio of approach 2. can be shown analogously.

Theorem 6.1
Let the value S_t of an object be a geometric Brownian motion (6.1). Let $\mathcal{U}$ be a derivative contingent on the object and maturing in T. Let $T^* \leq T$ be the time at which the derivative is exercised, or rather $T^* = T$ if it is not. Let the derivative's value at any time $t \leq T^*$ be given by a function $F(S_t, t)$ of the object's price and time.

a) *It exists a portfolio consisting of the underlying object and risk–less bonds which duplicates the derivative in the sense that it generates up to time T^* the same cash flow pattern as $\mathcal{U}$, and that it has the same time T^* value as $\mathcal{U}$.*

b) The derivatives value function $F(S,t)$ satisfies Black–Scholes partial differential equation

$$\frac{\partial F(S,t)}{\partial t} - rF(S,t) + bS\frac{\partial F(S,t)}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F(S,t)}{\partial S^2} = 0, \quad t \le T^*. \tag{6.3}$$

Proof:
To simplify we proceed from the assumption that the object is a stock paying a continuous dividend yield d, and thus involving costs of carry $b = r - d$ with r the continuous compounded risk–free interest rate. Furthermore, we consider only the case where $\mathcal{U}$ is a derivative on the stock, and that $\mathcal{U}$ does not generate any payoff before time T^*.

We construct a portfolio consisting of $n_t = n(S_t, t)$ shares of the stock and $m_t = m(S_t, t)$ zero bonds with maturity date T and a face value of $B_T = 1$. Let

$$B_t = B_T e^{-r(T-t)} = e^{-r(t-T)}$$

be the zero bond's value discounted to time t. We denote the time t portfolio value by

$$V_t \stackrel{\text{def}}{=} V(S_t, t) = n(S_t, t) \cdot S_t + m(S_t, t) \cdot B_t.$$

It is to show that n_t and m_t can be chosen such that at exercise time respectively at maturity of $\mathcal{U}$ the portfolio value is equal to the derivative's value, i.e. $V(S_{T^*}, T^*) = F(S_{T^*}, T^*)$. Furthermore, it is shown that the portfolio does not generate any cash flow prior to T^*, i.e. it is neither allowed to withdraw nor to add any money before time T^*. All changes in the positions must be realized by buying or selling stocks or bonds, or by means of dividend yields.

First of all, we investigate how the portfolio value V_t changes in a small period of time dt. By doing this, we use the notation $dV_t = V_{t+dt} - V_t,\ dn_t = n_{t+dt} - n_t$ etc.

$$\begin{aligned} dV_t &= n_{t+dt}S_{t+dt} + m_{t+dt}B_{t+dt} - n_t S_t - m_t B_t \\ &= dn_t S_{t+dt} + n_t dS_t + dm_t B_{t+dt} + m_t dB_t, \end{aligned}$$

and thus

$$dV_t = dn_t(S_t + dS_t) + n_t dS_t + dm_t(B_t + dB_t) + m_t dB_t. \tag{6.4}$$

Since the stochastic process S_t is a geometric Brownian motion and therefore an Itô–process (5.8) with $\mu(x,t) = \mu x$ and $\sigma(x,t) = \sigma x$, it follows from the generalized Itô lemma (5.10) and equation (6.1)

$$dn_t = \frac{\partial n_t}{\partial t}dt + \frac{\partial n_t}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 n_t}{\partial S^2}\sigma^2 S_t^2 dt, \tag{6.5}$$

and an analogous relation for m_t. Using

$$(dS_t)^2 = (\mu S_t dt + \sigma S_t dW_t)^2 = \sigma^2 S_t^2 (dW_t)^2 + o(dt) = \sigma^2 S_t^2 dt + o(dt),$$

$$dB_t = rB_t dt,\ dS_t \cdot dt = o(dt) \text{ and } dt^2 = o(dt)$$

and neglecting terms of size smaller than dt it follows:

$$dn_t(S_t + dS_t) = \Big(\frac{\partial n_t}{\partial t}dt + \frac{\partial n_t}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 n_t}{\partial S^2}\sigma^2 S_t^2 dt\Big)S_t + \frac{\partial n_t}{\partial S}\sigma^2 S_t^2 dt, \quad (6.6)$$

$$dm_t(B_t + dB_t) = \Big(\frac{\partial m_t}{\partial t}dt + \frac{\partial m_t}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 m_t}{\partial S^2}\sigma^2 S_t^2 dt\Big)B_t. \quad (6.7)$$

The fact that neither the derivative nor the duplicating portfolio generates any cash flow before time T^* means that the terms $dn_t(S_t+dS_t)$ and $dm_t(B_t+dB_t)$ of dV_t in equation (6.4) which correspond to purchases and sales of stocks and bonds have to be financed by the dividend yields. Since one share of the stock pays in a small time interval dt a dividend amount of $d \cdot S_t \cdot dt$, we have

$$d \cdot n_t S_t \cdot dt = (r - b) \cdot n_t S_t \cdot dt = dn_t(S_t + dS_t) + dm_t(B_t + dB_t).$$

Substituting equations (6.6) and (6.7) in the latter equation, it holds:

$$\begin{aligned} 0 \quad = \quad & (b-r)n_t S_t dt + \Big(\frac{\partial m_t}{\partial t}dt + \frac{\partial m_t}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 m_t}{\partial S^2}\sigma^2 S_t^2 dt\Big)B_t \\ & + \Big(\frac{\partial n_t}{\partial t}dt + \frac{\partial n_t}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 n_t}{\partial S^2}\sigma^2 S_t^2 dt\Big)S_t + \frac{\partial n_t}{\partial S}\sigma^2 S_t^2 dt. \end{aligned}$$

Using equation (6.1) and summarizing the stochastic terms with differential dW_t as well as the deterministic terms with differential dt containing the drift parameter μ, and all other deterministic terms gives:

$$\begin{aligned} 0 \quad = \quad & \Big(\frac{\partial n_t}{\partial S}S_t + \frac{\partial m_t}{\partial S}B_t\Big)\mu S_t dt \\ & + \Big\{\Big(\frac{\partial n_t}{\partial t} + \frac{1}{2}\frac{\partial^2 n_t}{\partial S^2}\sigma^2 S_t^2\Big)S_t + \frac{\partial n_t}{\partial S}\sigma^2 S_t^2 \\ & + \Big(\frac{\partial m_t}{\partial t} + \frac{1}{2}\frac{\partial^2 m_t}{\partial S^2}\sigma^2 S_t^2\Big)B_t + (b-r)n_t S_t\Big\}dt \\ & + \Big(\frac{\partial n_t}{\partial S}S_t + \frac{\partial m_t}{\partial S}B_t\Big)\sigma S_t dW_t. \end{aligned} \quad (6.8)$$

This is only possible if the stochastic terms disappear, i.e.

$$\frac{\partial n_t}{\partial S} S_t + \frac{\partial m_t}{\partial S} B_t = 0. \tag{6.9}$$

Thus the first term in (6.8) is neutralized as well. Hence the middle term must also be zero:

$$\begin{aligned} &\Big(\frac{\partial n_t}{\partial t} + \frac{1}{2}\frac{\partial^2 n_t}{\partial S^2}\sigma^2 S_t^2\Big) S_t + \frac{\partial n_t}{\partial S}\sigma^2 S_t^2 \\ +&\Big(\frac{\partial m_t}{\partial t} + \frac{1}{2}\frac{\partial^2 m_t}{\partial S^2}\sigma^2 S_t^2\Big) B_t + (b-r) n_t S_t = 0. \end{aligned} \tag{6.10}$$

To further simplify we compute the partial derivative of equation (6.9) with respect to S :

$$\frac{\partial^2 n_t}{\partial S^2} S_t + \frac{\partial n_t}{\partial S} + \frac{\partial^2 m_t}{\partial S^2} B_t = 0 \tag{6.11}$$

and substitute this in equation (6.10). We then obtain

$$\frac{\partial n_t}{\partial t} S_t + \frac{\partial m_t}{\partial t} B_t + \frac{1}{2}\frac{\partial n_t}{\partial S}\sigma^2 S_t^2 + (b-r) n_t S_t = 0. \tag{6.12}$$

Since the stock price S_t does not depend explicitly on time, i.e. $\partial S_t/\partial t = 0$, the derivative of the portfolio value $V_t = n_t S_t + m_t B_t$ with respect to time gives:

$$\frac{\partial V_t}{\partial t} = \frac{\partial n_t}{\partial t} S_t + \frac{\partial m_t}{\partial t} B_t + m_t \frac{\partial B_t}{\partial t} = \frac{\partial n_t}{\partial t} S_t + \frac{\partial m_t}{\partial t} B_t + m_t r B_t.$$

This implies

$$\frac{\partial n_t}{\partial t} S_t + \frac{\partial m_t}{\partial t} B_t = \frac{\partial V_t}{\partial t} - r m_t B_t = \frac{\partial V_t}{\partial t} - r(V_t - n_t S_t).$$

Substituting this equation in equation (6.12) we eliminate m_t and obtain

$$\frac{1}{2}\sigma^2 S_t^2 \frac{\partial n}{\partial S} + \frac{\partial V_t}{\partial t} + b n_t S_t - r V_t = 0. \tag{6.13}$$

Since the zero bond value B_t is independent of the stock price S_t, i.e. $\partial B_t/\partial S = 0$, the derivative of the portfolio value $V_t = n_t S_t + m_t B_t$ with respect to the stock price gives (using equation (6.9))

$$\frac{\partial V_t}{\partial S} = \frac{\partial n_t}{\partial S} S_t + n_t + \frac{\partial m_t}{\partial S} B_t = n_t,$$

and thus

$$n_t = \frac{\partial V_t}{\partial S}. \tag{6.14}$$

That is, n_t is equal to the so–called *delta* or hedge–ratio of the portfolio (see Section 6.3.1). Since

$$m_t = \frac{V_t - n_t S_t}{B_t}$$

we can construct a duplicating portfolio if we know $V_t = V(S_t, t)$. We can obtain this function of stock price and time as a solution of the Black–Scholes differential equation

$$\frac{\partial V(S,t)}{\partial t} - rV(S,t) + bS\frac{\partial V(S,t)}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(S,t)}{\partial S^2} = 0 \tag{6.15}$$

which results from substituting equation (6.14) in equation (6.13). To determine V we have to take into account a boundary condition which is obtained from the fact that the cash flows at exercising time respectively maturity, i.e. at time T^*, of the duplicating portfolio and the derivative are equal:

$$V(S_{T^*}, T^*) = F(S_{T^*}, T^*). \tag{6.16}$$

Since the derivative has at any time the same cash flow as the duplicating portfolio, $F(S,t)$ also satisfies the Black–Scholes differential equation, and at any time $t \le T^*$ it holds $F_t = F(S_t, t) = V(S_t, t) = V_t$. □

Black–Scholes' differential equation fundamentally relies on the assumption that the stock price can be modelled by a geometric Brownian motion. This assumption is justified, however, if the theory building on it reproduces the arbitrage relations derived in Chapter 2. Considering an example we verify this feature. Let $V(S_t, t)$ be the value of a future contract with delivery price K and maturity date T. The underlying object involves costs of carry at a continuous rate b. Since $V(S_t, t)$ depends only on the price of the underlying and time it satisfies the conditions of Theorem 6.1. From Theorem 2.1 and substituting $\tau = T - t$ for the time to maturity it follows

$$V(S,t) = Se^{(r-b)(t-T)} - Ke^{r(t-T)}.$$

Substituting the above equation into equation (6.3) it can be easily seen that it is the unique solution of Black–Scholes' differential equation with boundary condition $V(S,T) = S - K$. Hence, Black–Scholes' approach gives the same price for the future contract as the model free no–arbitrage approach.

Finally, we point out that modelling stock prices by geometric Brownian motion gives reasonable solutions for short and medium terms. Applying the model to other underlyings such as currencies or bonds is more difficult. Bond options typically have significant longer time to maturity than stock options. Their value does not only depend on the bond price but also on interest rates which have to be considered stochastic. Modelling interest rates reasonably involves other stochastic process, which we will discuss in later chapters.

Generally exchange rates cannot be modelled by geometric Brownian motion. Empirical studies show that the performance of this model depends on the currency and on the time to maturity. Hence, applying Black–Scholes' approach to currency options has to be verified in each case. If the model is used, the foreign currency has to be understood as the option underlying with a continuous foreign interest rate d corresponding to the continuous dividend yield of a stock. Thus, continuous costs of carry with rate $b = r - d$ equal the interest rate differential between the domestic and the foreign market. If the investor buys the foreign currency early, then he cannot invest his capital at home any more, and thus he looses the domestic interest rate r. However, he can invest his capital abroad and gain the foreign interest rate d. The value of the currency option results from solving Black–Scholes' differential equation (6.3) respecting the boundary condition implied by the option type.

6.2 Black–Scholes Formulae for European Options

In this section we are going to use Black–Scholes' equation to compute the price of European options. We keep the notation introduced in the previous chapter. That is, we denote

$$C(S,t) = C_{K,T}(S,t), \qquad P(S,t) = P_{K,T}(S,t)$$

the value of a European call respectively put option with exercise price K and maturity date T at time $t \leq T$, where the underlying, for example a stock, at time t has a value of $S_t = S$. The value of a call option thus satisfies for all prices S with $0 < S < \infty$ the differential equation

$$rC(S,t) - bS\frac{\partial C(S,t)}{\partial S} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C(S,t)}{\partial S^2} = \frac{\partial C(S,t)}{\partial t},\ 0 \leq t \leq T, \tag{6.17}$$

$$C(S,T) = \max\{0, S-K\},\ 0 < S < \infty, \tag{6.18}$$

$$C(0,t) = 0, \quad \lim_{S\to\infty} C(S,t) - S = 0,\ 0 \leq t \leq T. \tag{6.19}$$

The first boundary condition (6.18) follows directly from the definition of a call option, which will only be exercised if $S_T > K$ thereby procuring the gain $S_T - K$. The definition of Brownian motion implies that the process is absorbed by zero. In other words, if $S_t = 0$ for one $t < T$ it follows $S_T = 0$. That is the call will not be exercised, which is formulated in the first part of condition (6.19). Whereas the second part of (6.19) results from the reflection that the probability that the Brownian motion falls below K is fairly small if it attained a level significantly above the exercise price. If $S_t \gg K$ for a $t < T$ then it holds with a high probability that $S_T \gg K$. The call will be, thus, exercised and procures the cash flow $S_T - K \approx S_T$.

The differential equation (6.17) subject to boundary conditions (6.18). (6.19) can be solved analytically. To achieve this, we transform it into a differential equation known from the literature. First of all, we substitute the time variable t for the time to maturity $\tau = T - t$. By doing this, the problem with final condition (6.18) in $t = T$ changes to a problem subject to an initial condition in $\tau = 0$. Following, we multiply (6.17) by $2/\sigma^2$ and substitute the parameters r, b for

$$\alpha = \frac{2r}{\sigma^2},\ \beta = \frac{2b}{\sigma^2},$$

as well as the variables τ, S for

$$v = \sigma^2(\beta - 1)^2\frac{\tau}{2}, \qquad u = (\beta - 1)\ln\frac{S}{K} + v.$$

While for the original parameters hold $0 \le S < \infty, 0 \le t \le T$, for now the new parameters it holds

$$-\infty < u < \infty, \qquad 0 \le v \le \frac{1}{2}\sigma^2(\beta - 1)^2 T \stackrel{\text{def}}{=} v_T.$$

Finally, we set

$$g(u, v) = e^{r\tau} C(S, T - \tau)$$

and obtain the new differential equation

$$\frac{\partial^2 g(u, v)}{\partial u^2} = \frac{\partial g(u, v)}{\partial v}. \tag{6.20}$$

with the initial condition

$$g(u, 0) = K \max\{0, e^{\frac{u}{\beta-1}} - 1\} \stackrel{\text{def}}{=} g_0(u), \quad -\infty < u < \infty. \tag{6.21}$$

Problems with initial conditions of this kind are well known from the literature on partial differential equations. They appear, for example, in modelling heat conduction and diffusion processes. The solution is given by

$$g(u, v) = \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi v}} g_0(\xi) e^{-\frac{(\xi - u)^2}{4v}} d\xi.$$

The option price can be obtained by undoing the above variables and parameter substitutions. In the following we denote, as in Chapter 2, by $C(S,\tau)$ the call option price being a function of the time to maturity $\tau = T - t$ instead of time t. Then it holds

$$C(S,\tau) = e^{-r\tau} g(u,v) = e^{-r\tau} \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi v}} g_0(\xi) e^{-\frac{(\xi-u)^2}{4v}} d\xi.$$

Substituting $\xi = (\beta - 1)\ln(x/K)$ we obtain the original terminal condition $\max\{0, x-K\}$. Furthermore, replacing u and v by the variables S and τ we obtain

$$\begin{aligned} C(S,\tau) &= e^{-r\tau} \int_0^{\infty} \max(0, x-K) \frac{1}{\sqrt{2\pi}\sigma\sqrt{\tau}x} \\ &\quad \exp\left\{-\frac{[\ln x - \{\ln S + (b - \frac{1}{2}\sigma^2)\tau\}]^2}{2\sigma^2\tau}\right\} dx. \end{aligned} \tag{6.22}$$

In the case of Brownian motion $S_T - S_t$ is lognormally distributed, i.e. $\ln(S_T - S_t)$ is normally distributed with parameters $(b - \frac{1}{2}\sigma^2)\tau$ and $\sigma^2\tau$. The conditional distribution of S_T given $S_t = S$ is therefore lognormal as well but with parameters $\ln S + (b - \frac{1}{2}\sigma^2)\tau$ and $\sigma^2\tau$. However, the integrant in equation (6.22) is except for the term $\max(0, x-K)$ the density of the latter distribution. Thus, we can interpret the price of a call as the discounted expected option payoff $\max(0, S_T - K)$, which is the terminal condition, given the current stock price S :

$$C(S,\tau) = e^{-r\tau}\, \mathsf{E}[\max(0, S_T - K) \mid S_t = S]. \tag{6.23}$$

This property is useful when deriving numerical methods to compute option prices. But before doing that, we exploit the fact that equation (6.22) contains an integral with respect to the density of the lognormal distribution to further simplify the equation. By means of a suitable substitution we transform the term in an integral with respect to the density of the normal distribution and we obtain

$$C(S,\tau) = e^{(b-r)\tau} S\Phi(y + \sigma\sqrt{\tau}) - e^{-r\tau} K\Phi(y), \tag{6.24}$$

where we use y as a shortcut for

$$y = \frac{\ln\frac{S}{K} + (b - \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}. \tag{6.25}$$