Financial Statistics and Mathematical Finance
Methods, Models and Applications
Ansgar Steland, Institute for Statistics and Economics, RWTH Aachen University, Germany

Mathematical finance has grown into a huge area of research which requires a lot of care and a large number of sophisticated mathematical tools. Mathematically rigorous and yet accessible to advanced level practitioners and mathematicians alike, it considers various aspects of the application of statistical methods in finance and illustrates some of the many ways that statistical tools are used in financial applications.

Financial Statistics and Mathematical Finance:
• Provides an introduction to the basics of financial statistics and mathematical finance.
• Explains the use and importance of statistical methods in econometrics and financial engineering.
• Illustrates the importance of derivatives and calculus to aid understanding in methods and results.
• Looks at advanced topics such as martingale theory, stochastic processes and stochastic integration.
• Features examples throughout to illustrate applications in mathematical and statistical finance.
• Is supported by an accompanying website featuring R code and data sets.

Financial Statistics and Mathematical Finance introduces the financial methodology and the relevant mathematical tools in a style that is both mathematically rigorous and yet accessible to advanced level practitioners and mathematicians alike. Both graduate students and researchers in statistics, finance, econometrics and business administration will benefit from this book.
Financial Statistics and Mathematical Finance
Financial Statistics and Mathematical Finance

Methods, Models and Applications

Ansgar Steland

Institute for Statistics and Economics
RWTH Aachen University, Germany
# Contents

## Preface

xi

## Acknowledgements

xv

## 1 Elementary financial calculus

1.1 Motivating examples

1.2 Cashflows, interest rates, prices and returns

1.2.1 Bonds and the term structure of interest rates

1.2.2 Asset returns

1.2.3 Some basic models for asset prices

1.3 Elementary statistical analysis of returns

1.3.1 Measuring location

1.3.2 Measuring dispersion and risk

1.3.3 Measuring skewness and kurtosis

1.3.4 Estimation of the distribution

1.3.5 Testing for normality

1.4 Financial instruments

1.4.1 Contingent claims

1.4.2 Spot contracts and forwards

1.4.3 Futures contracts

1.4.4 Options

1.4.5 Barrier options

1.4.6 Financial engineering

1.5 A primer on option pricing

1.5.1 The no-arbitrage principle

1.5.2 Risk-neutral evaluation

1.5.3 Hedging and replication

1.5.4 Nonexistence of a risk-neutral measure

1.5.5 The Black–Scholes pricing formula

1.5.6 The Greeks

1.5.7 Calibration, implied volatility and the smile

1.5.8 Option prices and the risk-neutral density

1.6 Notes and further reading

## References

43

## 2 Arbitrage theory for the one-period model

2.1 Definitions and preliminaries

2.2 Linear pricing measures

2.3 More on arbitrage

45
# CONTENTS

2.4 Separation theorems in $\mathbb{R}^n$ .......................... 53
2.5 No-arbitrage and martingale measures .................. 56
2.6 Arbitrage-free pricing of contingent claims .......... 65
2.7 Construction of martingale measures: general case ... 70
2.8 Complete financial markets .............................. 73
2.9 Notes and further reading .................................. 76
References ....................................................... 76

3 Financial models in discrete time .......................... 79
3.1 Adapted stochastic processes in discrete time .......... 81
3.2 Martingales and martingale differences ................. 85
  3.2.1 The martingale transformation ...................... 91
  3.2.2 Stopping times, optional sampling and a maximal inequality 93
  3.2.3 Extensions to $\mathbb{R}^d$ ......................... 101
3.3 Stationarity .............................................. 102
  3.3.1 Weak and strict stationarity ...................... 102
3.4 Linear processes and ARMA models .................... 111
  3.4.1 Linear processes and the lag operator ............ 111
  3.4.2 Inversion ...................................... 116
  3.4.3 AR($p$) and AR($\infty$) processes ............. 119
  3.4.4 ARMA processes ................................ 122
3.5 The frequency domain .................................. 124
  3.5.1 The spectrum .................................. 124
  3.5.2 The periodogram ............................... 126
3.6 Estimation of ARMA processes .......................... 132
3.7 (G)ARCH models ...................................... 133
3.8 Long-memory series ................................... 139
  3.8.1 Fractional differences .......................... 139
  3.8.2 Fractionally integrated processes .............. 144
3.9 Notes and further reading ................................ 144
References ....................................................... 145

4 Arbitrage theory for the multiperiod model ............ 147
4.1 Definitions and preliminaries .......................... 148
4.2 Self-financing trading strategies ....................... 148
4.3 No-arbitrage and martingale measures ................. 152
4.4 European claims on arbitrage-free markets ............ 154
4.5 The martingale representation theorem in discrete time 159
4.6 The Cox–Ross–Rubinstein binomial model .............. 160
4.7 The Black–Scholes formula ............................. 165
4.8 American options and contingent claims ............... 171
  4.8.1 Arbitrage-free pricing and the optimal exercise strategy 171
  4.8.2 Pricing American options using binomial trees ... 174
4.9 Notes and further reading ................................ 175
References ....................................................... 175
## CONTENTS

8.4 Asymptotics 283  
8.5 Density estimation and nonparametric regression 287  
8.5.1 Multivariate density estimation 288  
8.5.2 Nonparametric regression 295  
8.6 The CLT for linear processes 302  
8.7 Mixing processes 306  
8.7.1 Mixing coefficients 306  
8.7.2 Inequalities 308  
8.8 Limit theorems for mixing processes 313  
8.9 Notes and further reading 323  
References 323

9 Special topics 325  
9.1 Copulas – and the 2008 financial crisis 325  
9.1.1 Copulas 326  
9.1.2 The financial crisis 332  
9.1.3 Models for credit defaults and CDOs 335  
9.2 Local Linear nonparametric regression 338  
9.2.1 Applications in finance: estimation of martingale measures and Itô diffusions 339  
9.2.2 Method and asymptotics 340  
9.3 Change-point detection and monitoring 350  
9.3.1 Offline detection 351  
9.3.2 Online detection 359  
9.4 Unit roots and random walk 363  
9.4.1 The OLS estimator in the stationary AR(1) model 364  
9.4.2 Nonparametric definitions for the degree of integration 368  
9.4.3 The Dickey–Fuller test 370  
9.4.4 Detecting unit roots and stationarity 373  
9.5 Notes and further reading 381  
References 382

Appendix A 385  
A.1 (Stochastic) Landau symbols 385  
A.2 Bochner’s lemma 387  
A.3 Conditional expectation 387  
A.4 Inequalities 388  
A.5 Random series 389  
A.6 Local martingales in discrete time 389

Appendix B Weak convergence and central limit theorems 391  
B.1 Convergence in distribution 391  
B.2 Weak convergence 392
Preface

This textbook intends to provide a careful and comprehensive introduction to some of the most important mathematical topics required for a thorough understanding of financial markets and the quantitative methods used there. For this reason, the book covers mathematical finance in the narrow sense, that is, arbitrage theory for pricing contingent claims such as options and the related mathematical machinery, as well as statistical models and methods to analyze data from financial markets. These areas evolved more or less separate from each other and the lack of material that covers both was a major motivation for me to work out the present textbook. Thus, I wrote a book that I would have liked when taking up the subject. It addresses master and Ph.D. students as well as researchers and practitioners interested in a comprehensive presentation of both areas, although many chapters can also be studied by Bachelor students who have passed introductory courses in probability calculus and statistics. Apart from a couple of exceptions, all results are proved in detail, although usually not in their most general form. Given the plethora of notions, concepts, models and methods and the resulting inherent complexity, particularly those coming to the subject for the first time can acquire a thorough understanding more quickly, if they can easily follow the derivations and calculations. For this reason, the mathematical formalism and notation is kept as elementary as possible. Each chapter closes with notes and comments on selected references, which may complement the presented material or are good starting points for further studies.

Chapter 1 starts with a basic introduction to important notions: financial instruments such as options and derivatives and related elementary methods. However, derivations are usually not given in order to focus on ideas, principles and basic results. It sets the scene for the following chapters and introduces the required financial slang. Cash flows, discounting and the term structure of interest rates are studied at an elementary level. The return over a given period of time, for assets usually a day, represents the most important economic object of interest in finance, as prices can be reconstructed from returns and investments are judged by comparing their return. Statistical measures for their location, dispersion and skewness have important economic interpretations, and the relevant statistical approaches to estimate them are carefully introduced. Measuring the risk associated with an investment requires being aware of the properties of related statistical estimates. For example, volatility is primarily related to the standard deviation and value-at-risk, by definition, requires the study of quantiles and their statistical estimation. The first chapter closes with a primer on option pricing, which introduces the most important notions of the field of mathematical finance in the narrow sense, namely the principle of no-arbitrage, the principle of risk-neutral pricing and the relation of those notions to probability calculus, particularly to the existence of an equivalent martingale measure. Indeed, these basic concepts and a couple of fundamental insights can be understood by studying them in the most elementary form or simply by examples.

Chapter 2 then discusses arbitrage theory and the pricing of contingent claims within a one-period model. At time 0 one sets up a portfolio and at time 1 we look at the result. Within this simple framework, the basic results discussed in Chapter 1 are treated with mathematical rigor and extended from a finite probability space, where only a finite number of scenarios
can occur, to a general underlying probability space that models the real financial market. Mathematical separation theorems, which tell us how one can separate a given point from convex sets, are applied in order to establish the equivalence of the exclusion of arbitrage opportunities and the existence of an equivalent martingale measure. For this reason, those separation theorems are explicitly proved. The construction of equivalent martingale measures based on the Esscher transform is discussed as well.

Chapter 3 provides a careful introduction to stochastic processes in discrete time (time series), covering martingales, martingale differences, linear processes, ARMA and GARCH processes as well as long-memory series. The notion of a martingale is fundamental for mathematical finance, as one of the key results asserts that in any financial market that excludes arbitrage, there exists a probability measure such that the discounted price series of a risky asset forms a martingale and the pricing of contingent claims can be done by risk-neutral pricing under that measure. These key insights allow us to apply the elaborated mathematical theory of martingales. However, the treatment in Chapter 3 is restricted to the most important findings of that theory, which are really used later. Taking first-order differences of a martingale leads naturally to martingale difference sequences, which form white-noise processes and are a common replacement for the unrealistic i.i.d. error terms in stochastic models for financial data and, more generally, economic data. A key empirical insight of the statistical analysis of financial return series is that they can often be assumed to be uncorrelated, but they are usually not independent. However, other series may exhibit substantial serial dependence that has to be taken into account. Appropriate parametric classes of time-series models are ARMA processes, which belong to the more general and infinite-dimensional class of linear processes. Basic approaches to estimate autocovariance functions and the parameters of ARMA models are discussed. Many financial series exhibit the phenomenon of conditional heteroscedasticity, which has given rise to the class of (G)ARCH models. Lastly, fractional differences and long-memory processes are introduced.

Chapter 4 discusses in detail arbitrage theory in a discrete-time multiperiod model. Here, trading is allowed at a finite number of time points and at each time point the trading strategy can be updated using all available information on market prices. Using the martingale theory in discrete time studied in Chapter 3, it allows us to investigate the pricing of options and other derivatives on arbitrage-free financial markets. The Cox–Ross–Rubinstein binomial model is studied in greater detail, since it is a standard tool in practice and also provides the basis to derive the famous Black–Scholes pricing formula for a European call. In addition, the pricing of American claims is studied, which requires some more advanced results from the theory of optimal stopping.

Chapter 5 introduces the reader to stochastic processes in continuous time. Brownian motion will be the random source that governs the price processes of our financial market model in continuous time. Nevertheless, to keep the chapter concise, the presentation of Brownian motion is limited to its definition and the most important properties. Brownian motion has puzzling properties such as continuous paths that are nowhere differentiable or of bounded variation. Advanced models also incorporate fractional Brownian motion and Lévy processes, respectively. Lévy processes inherit independent increments but allow for non-normal distributions of those increments including heavy tails and jump. Fractional Brownian motion is a Gaussian process as is Brownian motion, but it allows for long-range dependent increments where temporal correlations die out very slowly.

Chapter 6 treats the theory of stochastic integration. Assuming that the reader is familiar with integration in the sense of Riemann or Lebesgue, we start with a discussion of stochastic
Riemann–Stieltjes (RS) integrals, a straightforward generalization of the Riemann integral. The related calculus is relatively easy and provides a good preparation for the Itô integral. It is also worth mentioning that the stochastic RS-integral definitely suffices to study many issues arising in statistics. However, the problems arising in mathematical finance cannot be treated without the Itô integral. The key observation is that the change of the value of position $x(t)$ in a stock at time $t$ over the period $[t, t + \delta]$ is, of course, given by $x_t \delta P_t$, where $\delta P_t = P_{t+\delta} - P_t$. Aggregating those changes over $n$ successive time intervals $[i\delta, (i+1)\delta]$, $i = 0, \ldots, n-1$, in order to determine the terminal value, results in the sum $\sum_{i=0}^{n-1} x(i\delta) \delta P_{i\delta}$. Now ‘taking the limit $\delta \to 0$’ leads to an integral $\int x_s dP_s$ with respect to the stock price, which cannot be defined in the Stieltjes sense, if the stock price is not of bounded variation. Here the Itô integral comes into play. A rich class of processes are Itô processes and the famous Itô formula asserts that smooth functions of Itô processes again yield Itô processes, whose representation as an Itô process can be explicitly calculated. Further, ergodic diffusion processes as an important class of Itô processes are introduced as well as Euler’s numerical approximation scheme, which also provides the common basis for statistical estimation and inference of discretely sampled ergodic diffusions.

Chapter 7 presents the Black–Scholes model, the mathematically idealized model to price derivatives which is still the benchmark continuous-time model in practice. Here one may either invest in a risky stock or deposit money in a bank account that pays a fixed interest. The Itô calculus of Chapter 6 provides the theoretical basis to develop the mathematical arbitrage theory in continuous time. The classic Black–Scholes model assumes that the volatility of the stock price is constant with respect to time, which is too restrictive in practice. Thus, we briefly discuss the required changes when the volatility is time dependent but deterministic. Finally, the generalized Black–Scholes model allows the interest rate of the ‘risk-less’ instrument to be random as well as dependent on time, thus covering the realistic situation that money not invested in stocks is used to buy, for example, AAA-rated government bonds.

Chapter 8 studies the asymptotic limit theory for discrete-time processes as required to construct and investigate present-day methods for decision making; that is, procedures for estimation, inference as well as model checking, using financial data in the widest sense (returns, indexes, prices, risk measures, etc.). The limit theorems, partly presented along the way when needed to develop methodologies, cover laws of large numbers and central limit theorems for martingale differences, linear processes as well as mixing processes. The methods discussed in greater detail cover the multiple linear regression with stochastic regressors, nonparametric density estimation, nonparametric regression and the estimation of autocovariances and the long-run variance. Those statistical tools are ubiquitous in the analysis of financial data.

Chapter 9 discusses some selected topics. Copulas have become an important tool for modeling high-dimensional distributions with powerful as well as dangerous applications in the pricing of financial instruments related to credits and defaults. As a matter of fact, these played an unlucky role in the 2008 financial crisis when a simplistic pricing model was applied to large-scale pricing of credit default obligations. For this reason, some of the major developments leading to the crisis are briefly reviewed, revealing the inherent complexity of financial markets as well as the need for sophisticated mathematical models and their application. Local polynomial estimation is studied in greater detail, since it has important applications to many problems arising in finance such as the estimation of risk-neutral densities conditional volatility or discretely observed diffusion processes. The asymptotic normality can be based on a powerful reduction principle: A (joint) smoothing central limit theorem for the innovation process $\{e_t\}$ and a derived process involving the regressors automatically
implies the asymptotic normality of the local linear estimator. The testing for and detecting of change-points (structural breaks) have become active fields of current theoretical as well as applied research. Chapter 9 thus closes with a brief discussion of change-point analysis and detection with a certain focus on the detection of changes in the degree of integration.

This book features an accompanying website http://fsmf.stochastik.rwth-aachen.de

Ansgar Steland
RWTH Aachen University
Germany
Acknowledgements

I thank the following students and assistants who carefully read parts of various drafts leading up to this book: Martin Euskirchen, who also typed some chapters, Dr. Wolfgang Herff and M. Sc. Annabel Prause. In particular, I acknowledge the assistance of Dr. Mohammed A bujarad, who proof-read the final version with great care and found many typos, mistakes and made several suggestions concerning the presentation of results. Lastly, I acknowledge the patience of my family and especially my children.
1

Elementary financial calculus

1.1 Motivating examples

Example 1.1.1 Suppose a pension fund collecting contributions from workers intends to invest a certain fraction of the fund in a certain exchange-traded stock instead of buying treasury bonds. Whereas a bond yields a fixed interest known in advance, the return of a stock is volatile and uncertain. It may substantially exceed a bond’s interest, but the pension fund is also exposed to the downside risk that the stock price goes down resulting in a loss. For the pension fund it is important to know what return can be expected from the investment and which risk is associated with the investment. It would also be useful to know the amount of the invested money that is under risk. In practice, investors invest their money in a portfolio of risky assets. Then the question arises: what can be said about the relationship? In modern finance, returns are modeled by random variables that have a distribution. Thus, we have to clarify how the return distribution and its mathematical properties are related to the economic notions expected return, volatility, and how one can define appropriate risk measures. Further, the question arises how one can estimate these quantities from historic time series.

Example 1.1.2 In order to limit the loss due to the risky stock investment, the pension fund could ask a bank for a contract that pays the difference between a stop loss quote, $L$, and stock price, if that difference is positive when exercising the contract. Such financial instruments are called options. What is the fair price of such an option? And how can a bank initiate trades, which compensate for the risk exposure when selling the option?

Example 1.1.3 Suppose a steel producer agrees with a car manufacturer to deliver steel for the production of 10 000 cars in one year. The steel production starts in one year and requires a large amount of oil. In order to calculate costs, the producer wants to fix the oil price at, say, $K$ dollars in advance. One approach is to enter a contract that pays the difference between the oil price and $K$ at the delivery date, if that difference is positive. Such contracts are named
call options. Again, the question arises what is the fair price of such an agreement. Another possibility is to agree on a future/forward contract.

**Example 1.1.4** To be more specific and to simplify the exposition, let us assume that the steel producer needs 1 barrel whose current price at time \( t = 0 \) is \( S_0 = 100 \). To fix that price, he buys a call option with delivery price \( K = 100 \). The fixed interest rate is 1%. Further, suppose that the oil price, \( S_1 \), in one year at time \( t = 1 \) is distributed according to a two-point distribution,

\[
P(S_1 = 110) = 0.6, \quad P(S_1 = 90) = 0.4.
\]

If \( S_1 = 110 \) one exercises the option right and the deal yields a profit of \( G = 10 \). Otherwise, the option has no value. Thus, the expected profit is given by

\[
E(G) = 10 \cdot 0.6 = 6.
\]

Because for the buyer of the option the deal has a non-negative profit and yields a positive profit with positive probability, he or she has to pay a premium to the bank selling the option. Should the bank offer the option for the expected profit 6? Surprisingly, the answer is no. Indeed, an oil dealer can offer the option for a lower price, namely \( x = 5.45 \) without making a loss. The dealer buys half of the oil when entering the contract at \( t = 0 \) for the current price of 50 and the rest when the contract is settled. His calculation is as follows. He finances the deal by the premium \( x \) and a credit. At \( t = 0 \) his portfolio consists of a position in the money market, \( x - 50 \), and 0.5 units of oil. Let us anticipate that \( x < 50 \). Then at \( t = 1 \) the dealer has to pay back \( 1.01 \cdot |x - 50| \) to the bank. We shall now consider separately the cases of an increase or decreases of the oil price. If the oil price increases, the value of the oil increases to \( 0.5 \cdot 110 = 55 \) and he receives 100 from the steel producer. He has to fix the premium \( x \) such that the net income equals the price he has to pay for the remaining oil. This means, he solves the equation

\[
100 + 1.01 \cdot (x - 50) = 55
\]

yielding \( x = 5.445545 \approx 5.45 \). Now consider the case that the oil price decreases to 90. In this case the steel producer does not exercise the option but buys the oil at the spot market. The oil dealer has to pay back the credit, sells his oil at the lower price, which results in a loss of 5. The premium \( x \) should ensure that his net balance is 0. This means, the equation

\[
0.5 \cdot 90 + 1.01(x - 50) = 0
\]

should hold. Solving for \( x \) again yields \( x = 5.445545 \). Notice that both equations yield the same solution \( x \) such that the premium is not random.

### 1.2 Cashflows, interest rates, prices and returns

Let us now introduce some basic notions and formulas. To any financial investment initiated at \( t = t_0 \) with time horizon \( T \) is attached a sequence of payments settled on a bank account that describe the investment from a mathematical point of view. Our standard notation is as follows: We denote the time points of the payments by \( 0 = t_0 < t_1 < \cdots < t_n = T \) and the associated payments by \( X_1, \ldots, X_T \). Our sign convention will be as follows: Positive payments, \( X_i > 0 \),
are deposits increasing the investor’s bank account, whereas negative payments, $X_i < 0$, are charges.

From an economic point of view, there is a huge difference between a payment today or in the future. Thus, to compare payments, they either have to refer to the same time point $t^*$ or one has to take into account the effects of interest. As a result, to compare investments one has to cumulate the payments discounted or accumulated to a common time point $t^*$. If all payments are discounted to $t^* = t_0$ and then cumulated, the resulting quantity is called the **present value**. Alternatively, one can accumulate all payments to $t^* = T$.

In practice, one has to specify how to determine times and how to measure the economic distance between two time points $t_1$ and $t_2$. It is common practice to measure the time as a multiple of a year. At this point, suppose that the dates are given using the day-month-year convention, i.e. $t = (d, m, y)$. In what follows, we denote the economic time distance between two dates $t_1$ and $t_2$ by $\tau(t_1, t_2)$. Here are some market conventions for the calculation of $\tau(t_1, t_2)$.

(i) Actual/365: Each year has 365 days and the actual number of days is used.

(ii) Actual/360: Each year has 360 days and the actual number of days is used.

(iii) 30/360: Each month has 30 days, a year 360 days.

In the following we assume that all times have been transformed using such a convention.

If the fixed interest rate is $r$ per annum, interest is paid during the period without compound interest, the accumulated value of payments $X_1, \ldots, X_n$ at dates $t_1, \ldots, t_n$ is given by

$$V_T = \sum_{i=0}^{n} X_i (1 + \tau(t_i, 1)r).$$

The present value at $t = 0$ is calculated using the formula

$$V_0 = \sum_{i=0}^{n} X_i D(0, t_i), \quad \text{with} \quad D(0, t_i) = \frac{1 + \tau(t_i, T)r}{1 + rT}.$$ 

Here $D(0, t_i)$ denotes the discount factor taking into account that the payment $X_i$ takes place at $t_i$.

Often, interest is paid at certain equidistant time points, e.g. quarterly or monthly. When decomposing the year into $m$ periods and applying the interest rate $r/m$ to each of them, an investment of one unit of currency grows during $k$ periods to

$$1 + \frac{r}{m}k.$$ 

When compound interest is taken into account, the value is

$$(1 + r/m)^k.$$ 

For $k = m \to \infty$ that discrete interest converges to continuous compounding

$$\lim_{m \to \infty} (1 + r/m)^m = e^r.$$
Thus, the accumulation factor for an investment lasting for \( t \in (0, \infty) \) years, i.e. corresponding to \( tm \) periods, equals

\[
\lim_{m \to \infty} (1 + r/m)^{mt} = e^{rt}.
\]

Let us now assume that the interest rate \( r = r(t) \) is a function of \( t \), such that for \( r(t) > 0 \), \( t > 0 \), the bank account, \( S_0(t) \), increases continuously. There are two approaches to relate these quantities. Either start from a model or formula for \( S_0(t) \) or start with \( r(t) \). Let us first suppose that \( S_0(t) \) is given. The annualized relative growth during the time interval \([t, t+h]\) is given by

\[
\frac{1}{h} \frac{S_0(t+h) - S_0(t)}{S_0(t)}.
\]

**Definition 1.2.1** Suppose that the bank account \( S_0(t) \) is a differentiable function. Then the quantity

\[
 r(t) = \lim_{h \to 0} \frac{1}{h} \frac{S_0(t+h) - S_0(t)}{S_0(t)},
\]

is well defined and is called **instantaneous (spot) rate** or simply **short rate**.

We have the relationship

\[
 r(t) = \frac{S_0'(t)}{S_0(t)} \iff S_0'(t) = r(t)S_0(t).
\]

As a differential:

\[
 dB(t) = r(t)B(t)dt.
\]

It is known that this ordinary differential equation has the general solution \( S_0(t) = C \exp(\int_0^t r(s) \, ds) \), \( C \in \mathbb{R} \). For our example the special solution

\[
 S_0(t) = \exp \left( \int_0^t r(s) \, ds \right) \tag{1.1}
\]

with starting value \( S_0(0) = 1 \) matters. In the special case \( r(t) = r \) for all \( t \), we obtain \( S_0(t) = e^{rt} \) as above.

Often, one starts with a model for the short rate. Then we define the bank account via Equation (1.1).

**Definition 1.2.2** (**Bank account**)  
A bank account with a unit deposit and continuous compounding according to the spot rate \( r(t) \) is given by

\[
 S_0(t) = \exp \left( \int_0^t r(s) \, ds \right), \quad t \geq 0.
\]

When depositing \( x \) units of currency into the bank account, the time \( t \) value is \( xS_0(t) \). Vice versa, for an accumulated value of 1 unit of currency at time \( T \), one has to deposit \( x = 1/S_0(T) \).
at time $t = 0$. The value of $x = 1/S_0(T)$ at an arbitrary time point $t \in [0, T]$ is
\[ xS_0(t) = \frac{S_0(t)}{S_0(T)}. \]
This means that the value at time $t = 0$ of a unit payment at the time horizon $T$ is given by $S_0(t)/S_0(T)$.

**Definition 1.2.3** The **discount factor** between two time points $t \leq T$ is the amount at time $t$ that is equivalent to a unit payment at time $T$ and can be invested riskless at the bank. It is denoted by
\[ D(t, T) = \frac{S_0(t)}{S_0(T)} = \exp\left(- \int_t^T r(s) \, ds \right). \]

### 1.2.1 Bonds and the term structure of interest rates

The basic insights of the above discussion can be directly used to price bonds and understand the term structure of interest rates.

A **zero coupon bond** pays a fixed amount of money, the **face value** or **principal** $X$ at a fixed future time point called **maturity**. Such a bond is also referred to as a **discount bond** or **zero coupon bond**. Here and in what follows, we assume that the bond is issued by a government such that we can ignore default risk. Measuring time in years and assuming that the interest rate $r$ applies in each year, we have learned that the present value of the payment $X$ equals
\[ P_n(X) = \frac{X}{(1 + r)^n}. \]

Notice that this simple formula determines a 1-to-1 correspondence between the bond price and the interest rate. The interest rate $r$ is the **discount rate** or **spot interest rate** for time to maturity $n$; spot rate, since that rate applies to a contract agreed on today.

Let us now consider a coupon bearing bond that pays coupons $C_1, \ldots, C_k$ at times $t_1, \ldots, t_k$ and the face value $X$ at the maturity date $T$. This series of payments is equivalent to $k + 1$ zero coupon bonds with face values $C_1, \ldots, C_k, X$ and maturity dates $t_1, \ldots, t_k, T$. Thus, its price is given by the **bond price equation**
\[ P(t) = \sum_{i=1}^k C_i P(t, t_i) + XP(t, T), \]
or equivalently
\[ P(t) = \sum_{i=1}^k C_i P(t, t + \tau_i) + XP(t, T), \]
if $\tau_j = t_j - t$ denotes the time to maturity of the $j$th bond. It follows that the price of the bond can be determined by the curve $\tau \mapsto P(t, t + \tau)$ that assigns to each maturity $\tau$ the time $t$ price for a zero coupon bond with unit principal $t$. It is called the **term structure of interest rates**.
There is a second approach to describe the term structure of interest rates. Let \( P(t, t + m) \) denote the price at time \( t \) of a zero coupon bond paying the principal \( X = 1 \) at the maturity date \( t + m \). Given the yearly spot rate \( r(t, t + m) \) applying to a payment in \( m \) years, its price is given by

\[
P(t, t + m) = \frac{1}{(1 + r(t, t + m))^m}.
\]

If the coupon corresponding to the interest rate \( r(t, t + m) \) is paid at \( n \) equidistant time points with continuous compounding, the formula

\[
P(t, t + m) = \frac{1}{(1 + r(t, t + m)/n)^{nm}}
\]

applies, which converges to the formula for continuously compounding

\[
P(t, t + m) = e^{-r(t,t+m)m} \quad \Leftrightarrow \quad P(t, T) = e^{-r(t,T)(T-t)},
\]

using the substitution \( T = t + m \). The continuously compounded interest rate \( r(t, T) \) is also called yield and the function

\[
t \mapsto r(t, T)
\]

the yield curve.

Finally, one can also capture the term structure of interest rates by the instantaneous forward rate at time \( t \) for the maturity date \( T \) defined by

\[
f(t, T) = \frac{\partial}{\partial T} P(t, T) = -\frac{\partial}{\partial T} \log P(t, T).
\]

Here it is assumed that the bond price \( P(t, T) \) is differentiable with respect to maturity. It then follows that

\[
P(t, T) = \exp \left( - \int_0^\tau f(t, t + s) \, ds \right),
\]

\[
r(t, t + \tau) = -\frac{1}{\tau} \int_0^\tau f(t, t + s) \, ds.
\]

### 1.2.2 Asset returns

For fixed-income investments such as treasury bonds the value of the investment can be calculated in advance, since the interest rate is known. By contrast, for assets such as exchange-traded stocks the interest rates, i.e., returns, are calculated from the quotes that reflect the market prices.

Let \( S_t \) be the price of a stock at time \( t \). Since such prices are quoted at certain (equidistant) time points, it is common to agree that the time index attains values in the discrete set of natural numbers, \( \mathbb{N} \). If an investor holds one share of the stock during the time interval from time \( t - 1 \) to \( t \), the asset price changes to

\[
S_t = S_{t-1}(1 + R_t),
\]

where

\[
R_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1.
\]
is called the simple net return and

$$1 + R_t = \frac{S_t}{S_{t-1}}$$

are the gross returns. How are asset returns aggregated over time? Suppose an investor holds a share between $s$ and $t = s + k$, i.e. over $k$ periods, $s, t, k \in \mathbb{N}$ (or more generally $s, t, k \in [0, \infty)$). Define the $k$-period return

$$R_t(k) = \frac{S_t - S_s}{S_s} = \frac{S_t}{S_s} - 1.$$

One easily checks the following relationship between the simple returns $R_{s+1}, \ldots, R_t$ and the $k$-period return:

$$1 + R_t(k) = \frac{S_t}{S_s} = \prod_{i=s+1}^{t} \frac{S_i}{S_{i-1}} = \prod_{i=s+1}^{t} (1 + R_i).$$

When an asset is held for $k$ years, the annualized average return (effective return) is given by the geometric mean

$$R_{t,k} = \left[ \prod_{i=0}^{k-1} (1 + R_{t+i}) \right]^{1/k} - 1.$$

A fixed-income investment with an annualized interest rate of $R_{t,k}$ yields the same accumulated value. Note that

$$R_{t,k} = \exp \left[ \frac{1}{k} \sum_{i=0}^{k-1} \log(1 + R_{t+i}) \right] - 1. \quad (1.2)$$

The natural logarithm of the gross returns,

$$r_t = \log(1 + R_t) = \log \frac{S_t}{S_{t-1}}$$

is called log return. Using Equation (1.2) we see that the $k$-period log return for the period from $s$ to $t = s + k$ can be calculated as

$$r_t(k) = \log(1 + R_t(k)) = \sum_{i=s+1}^{t} \log(1 + R_i) = \sum_{i=s+1}^{t} r_i.$$

Thus, in contrast to the returns $R_t$ the log returns possess the pleasant property of additivity w.r.t. time aggregation.

Using these definitions we obtain the following fundamental multiplicative decomposition of an asset price:

$$S_t = S_0 \prod_{i=1}^{t} (1 + R_i) = S_0 \prod_{i=1}^{t} \exp(r_i).$$
1.2.3 Some basic models for asset prices

When a security is listed on a stock exchange, there exists no quote before that time. Let us denote the sequence of price quotes, often the daily closing prices, by \( S_0, S_1, \ldots \). Since \( S_0 > 0 \) denotes the first quote, it is often regarded as a constant. If one wants to avoid possible effects of the initial price, one puts formally \( S_0 = 0 \).

A first approach for a stochastic model is to assume that the price differences are given by

\[ \Delta + u_n, \quad n = 1, 2, \ldots \]

with a deterministic, i.e. nonrandom, constant \( \Delta \in \mathbb{R} \) and i.i.d. random variables \( u_n, n \in \mathbb{N} \), with common distribution function \( F \) such that

\[ E(u_n) = 0, \quad \text{Var}(u_n) = \sigma^2 \in (0, \infty), \quad \forall n \in \mathbb{N}. \]

In the present context, it is common to name the \( u_n \) innovations. When referring to the sequence of innovations, we shall frequently write \( \{u_n : n \in \mathbb{N}_0\} \) or, for brevity of notation, \( \{u_n\} \) if the index set is obvious. The above model for the differences implies that the price process is given by

\[ S_t = S_0 + \sum_{i=1}^{t} (\Delta + u_i) = S_0 + t \Delta + \sum_{i=1}^{t} u_i, \quad t = 0, 1, \ldots \]

where we put \( u_0 = 0 \) and agree on the convention that \( \sum_{i=1}^{0} a_i = 0 \) for any sequence \( \{a_n\} \).

\( S_t \) is called (arithmetic) random walk and random walk with drift if \( \Delta \neq 0 \). Obviously

\[ E(S_t) = S_0 + \Delta t \]

and

\[ \text{Var}(S_t) = t\sigma^2. \]

This particular model for an asset price dates back to the work of Bachelier (1900).

An alternative approach is based on the log returns. Let us denote

\[ R_i := \log(S_i/S_{i-1}), \quad i \geq 1. \]

Then

\[ S_t = S_0 \prod_{i=1}^{t} S_i/S_{i-1} = S_0 \prod_{i=1}^{t} \exp(R_i). \]

The associated log price process is then given by

\[ \log S_t = \log S_0 + \sum_{i=1}^{t} R_i, \quad t = 0, 1, \ldots, \]

which is again a random walk.

A classic distributional assumption for the log returns \( \{R_n\} \) is the normal one,

\[ R_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2) \]
with \( \mu \in \mathbb{R} \) and \( \sigma^2 > 0 \). As a consequence, the log prices are normally distributed as well,

\[
\log(S_t) = \log(S_0) + \sum_{i=1}^t R_i \sim N(\log(S_0) + t\mu, t\sigma^2).
\]

Thus, \( S_t \) follows a lognormal distribution. Let us summarize some basic facts about that distribution:

A random variable \( X \) follows a lognormal distribution with parameters \( \mu \in \mathbb{R} \) (drift) and \( \sigma > 0 \) (volatility) if \( Y = \log(X) \sim N(\mu, \sigma^2) \). \( X \) takes on values in the interval \((0, \infty)\)

\[
P(\log X \leq y) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{y} e^{-(t-\mu)^2/2\sigma^2} dt, \quad y \in (0, \infty).
\]

The change of variable \( u = e^{t} \) leads to

\[
P(X \leq e^{y}) = P(\log X \leq y) = \int_{-\infty}^{e^{y}} \frac{1}{\sqrt{2\pi}\sigma u} e^{-(\log u-\mu)^2/2\sigma^2} du.
\]

By evaluating the right-hand side at \( y = \log x \), we see that the density \( f(x) \) of \( X \) is given by

\[
f(x) = \frac{1}{\sqrt{2\pi}x\sigma} e^{-(\log x-\mu)^2/2\sigma^2} 1(x > 0), \quad x \in \mathbb{R}.
\] (1.3)

Now it is easy to verify that mean and variance of \( X \) are given by

\[
E(X) = e^{\mu+\sigma^2/2} \quad \text{and} \quad \text{Var}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1).
\]

In order to model distributions that put more mass to extreme values than the standard normal distribution, one often uses the \( t \)-distribution with \( n \) degrees of freedom defined via the density function

\[
f(x) = \frac{1}{n\pi} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \left( 1 + \frac{x^2}{n} \right)^{-\frac{n+1}{2}}, \quad x \in \mathbb{R},
\]

which is parametrized by \( n \in \mathbb{N} \). By symmetry, its expectation is zero and the variance turns out to be \( n/(n-2) \), if \( n > 2 \).

Several questions arise: Which of the above two models holds true or provides a better approximation to reality? Are returns and log returns, respectively, normally distributed? Are asset returns symmetrically distributed? How can we estimate important distributional parameters such as \( \mu \), \( \sigma^2 \) or the skewness? Does the assumption of independent returns apply to real returns? Do price processes follow random walk models at all? What is the effect of changes of economic conditions on the distribution of returns? Can we test or detect such effects? How can we model the stochastic relationship between the return series of, say, \( m \) securities?

There is some evidence that some financial variables have much heavier tails than a normal distribution.
A random variable $X$ has a stable distribution or is stable, if $X$ has a domain of attraction. The latter means that there exist i.i.d. random variables $\{\xi_n\}$ and sequences $\{\sigma_n\} \subset (0, \infty)$ and $\{\mu_n\} \subset \mathbb{R}$, such that

$$ \frac{1}{\sigma_n} \sum_{i=1}^n \xi_i + \mu_n \xrightarrow{d} X, $$

as $n \to \infty$. The classic central limit theorem tells us that the $X \sim N(\mu, \sigma^2)$ is stable. By the Lévy–Khintchine formula, the characteristic function

$$ \phi(\theta) = E(e^{i\theta X}), \quad \theta \in \mathbb{R}, $$

where $i^2 = -1$, of a stable random variable $X$ has the representation

$$ \phi(\theta) = \begin{cases} 
\exp \left\{ i\mu \theta - \sigma^\alpha |\theta|^\alpha \left( 1 - i\beta \left( \text{sgn}(\theta) \right) \tan \frac{\pi \alpha}{2} \right) \right\}, & \alpha \neq 1, \\
\exp \left\{ i\mu \theta - \sigma |\theta| \left( 1 + i\beta \frac{2}{\pi} (\text{sgn}(\theta)) \log |\theta| \right) \right\}, & \alpha = 1,
\end{cases} $$

where $0 < \alpha \leq 2$ is the stability (characteristic) exponent, $-1 < \beta < 1$ the skewness parameter, $\sigma > 0$ the scale parameter and $\mu \in \mathbb{R}$ the location parameter. For $\alpha = 2$ one obtains the normal distribution $N(\mu, \sigma^2)$, since then $\phi(\theta) = \exp(i\mu \theta - \sigma^2 \theta^2)$. The tails of a standard normal distribution decay exponentially fast,

$$ P(|X| > x) \sim \sqrt{\frac{2}{\pi}} e^{-x^2/2}, \quad x \to \infty \quad (X \sim N(0, 1)). $$

By contrast, the tails of a stable random variable $X$ with characteristic exponent $0 < \alpha < 2$ decay as $x^{-\alpha}$, since

$$ \lim_{x \to \infty} x^\alpha P(X > x) = C_\alpha \frac{1 + \beta}{2} \sigma^\alpha $$

and

$$ \lim_{x \to \infty} x^\alpha P(X < -x) = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha, $$

where $C_\alpha = \left( \int_0^\infty x^{-\alpha} \sin(x) \, dx \right)^{-1}$.

Stable distributions appear as a special case of infinitely divisible distributions. A random variable (or random vector) $X$ and its distribution are called infinitely divisible, if for every $n \in \mathbb{N}$ there exist independent and identically distributed random variables $X_{n1}, \ldots, X_{nn}$ such that

$$ X \overset{d}{=} X_{n1} + \cdots + X_{nn}. $$

Those infinitely divisible distributions are exactly the distributions that can appear as limits of the distributions of sums $\sum_{k=1}^n X_{nk}$ of such arrays of row-wise i.i.d. random variables. Let $X$ be a $d$-dimensional random vector and again let $\phi(\theta) = E(\exp(i\theta' X))$, $\theta \in \mathbb{R}^d$, be its characteristic function. Then, the Lévy–Khintchine formula asserts that

$$ \phi(\theta) = \exp \left\{ i\theta' b - \frac{1}{2} \theta' C \theta + \int_{\mathbb{R}^d} \left( e^{i\theta' x} - 1 - i\theta' h(x) \right) \, dv(x) \right\}, $$

(1.6)
where
\[ h(x) = x1(|x| \leq 1), \quad x \in \mathbb{R}^d. \]
is a truncation function, \( b \in \mathbb{R}^d \) and \( C \) a symmetric and non-negative definite \((d \times d)\)-matrix and \( \nu \) a Lévy measure, that is a positive measure on the Borel sets of \( \mathbb{R}^d \) such that \( \nu(\{0\}) = 0 \) and
\[ \int_{\mathbb{R}^d} (|x|^2 \wedge 1) \, d\nu(x) < \infty. \]
As a consequence, \( \varphi(\theta) \) is characterized by the triplet \((b, C, \nu)\).

The characteristics of the normal distribution \( N(\mu, \sigma^2) \) are \((b, C, \nu) = (\mu, \sigma^2, 0)\), of course. For a Poisson distribution with intensity \( \lambda \), the characteristic function is
\[ \varphi(\theta) = \exp(\lambda(e^{i\theta} - 1)), \]
which results, if we put \( b = \lambda, C = 0 \) and \( \nu \) the one-point measure that assigns mass \( \lambda \) to the single point 1.

### 1.3 Elementary statistical analysis of returns

We have seen that price processes can be built from returns \( R_t \) that are modeled as random variables. For simplicity of our exposition, let us assume that \( R_1, \ldots, R_T \) are independent and identically distributed. To simplify notation, let \( R \) denote a generic return, i.e. \( R \overset{d}{=} R_1 \) which means that for any event \( A \) we have \( P(R \in A) = P(R_1 \in A) \).

But before focusing on returns, let us briefly review the most basic probabilistic quantities to which we will refer frequently in the following for an arbitrary random variable \( X \). In general, the distribution of a random variable is uniquely determined by its distribution function (d.f.)
\[ F_X(x) = P(X \leq x), \quad x \in \mathbb{R}. \]
If \( f : \mathbb{R} \to \mathbb{R} \) is a density, i.e. non-negative function with \( \int f(x) \, dx = 1 \), then the d.f. \( F(x) \) can be calculated by
\[ F(x) = \int_{-\infty}^{x} f(t) \, dt, \quad x \in \mathbb{R}. \]
A random variable \( X \) that attains a density function \( f \) is called a continuous random variable. Usually, it is assumed that returns are continuous random variables in that sense.

The first moment is defined by \( \mu = E(X) \) and can be calculated for a continuous random variable via
\[ \mu = E(X) = \int_{-\infty}^{\infty} xf(x) \, dx. \]
\( E(X) \) is also called the expectation or mean of \( X \). If \( X \) is a discrete random variable, that is \( X \) takes values in some discrete set \( \{x_1, x_2, \ldots\} \) of possible values with corresponding
probabilities \( p_1, p_2, \ldots \) such that
\[
P(X = x_i) = p_i, \quad i = 1, 2, \ldots
\]
then
\[
E(X) = \sum_{i=1}^{\infty} x_i p_i.
\]

More generally, the \( k \)th moment of \( X \) is defined as \( E(X^k) \) and \( E|X|^k \) is referred to as the \( k \)th absolute moment. Assumptions on the existence of higher moments control the probability of outliers, that is extreme values. Indeed, by virtue of Markov’s inequality, the probability that \( X \) takes values larger than \( c > 0 \) in absolute value decays faster for increasing \( c \), if higher moments exist, since
\[
P(|X| > c) \leq \frac{E|X|^k}{c^k}.
\]

Compare this inequality with the formulas (1.4) and (1.5) for the special class of stable distributions. As extreme values (outliers) of daily returns, usually negative ones, correspond to unexpected high-impact news such as a crash, the behavior of the tail probabilities \( P(X < -c) \) and \( P(X > c) \), \( c > 0 \), are of substantial interest, and moment assumptions automatically constrain them.

Suppose we are given a random sample \( X_1, \ldots, X_T \) of sample size \( T \). The empirical distribution function of the sample \( X_1, \ldots, X_T \) is defined as
\[
F_T(x) = \frac{1}{T} \sum_{t=1}^{T} 1(X_t \leq x), \quad x \in \mathbb{R}.
\]
Notice that \( F_T(x) \) is the fraction of observations that are less or equal than \( x \).

For a distribution function \( F \) let
\[
F^{-1}(y) = \inf \{x : F(x) \geq y\}
\]
denote the left-continuous inverse called quantile function. Applying that definition to the empirical distribution function yields the sample quantile function
\[
F_T^{-1}(p) = \inf \{x : F_T(x) \geq p\} = X_{\lfloor np \rfloor}, \quad p \in (0, 1).
\]
For a fixed \( p \), \( F_T^{-1}(p) \) is called the sample \( p \)-quantile or empirical \( p \)-quantile. Here \( X_{(1)} \leq \cdots \leq X_{(T)} \) denotes the order statistic and \( \lfloor x \rfloor \) is the smallest integer larger or equal to \( x \). Notice that \( X_{\lfloor np \rfloor} = X_{\lfloor np \rfloor + 1} \) where \( \lfloor x \rfloor \) is the floor function, i.e. the largest integer that is less than or equal to \( x \). Quantiles play an important role in characterizing a distribution. The sample 0.5-quantile is called the median and is also denoted by \( x_{\text{med}} \). Together with the 0.25- and 0.75-quantiles,
\[
Q_1 = F_T^{-1}(0.25), \quad Q_3 = F_T^{-1}(0.75),
\]
called quartiles, we get a picture where the lower (upper) fourth and the central 50% of the data are located. Augmenting these three statistics with the minimum and maximum defining the