Universitext

UTX

Szymon Borak
Wolfgang Karl Härdle
Brenda López-Cabrera

Statistics of Financial Markets

Exercises and Solutions

Second Edition

NOW WITH EXOTIC OPTIONS
AND MORE QUANTLETS

R and Matlab Quantlets

QUANTLETS

Springer

Universitext

Universitext

Universitext is a series of textbooks that presents material from a wide variety of mathematical disciplines at master's level and beyond. The books, often well class-tested by their author, may have an informal, personal even experimental approach to their subject matter. Some of the most successful and established books in the series have evolved through several editions, always following the evolution of teaching curricula, to very polished texts.

Thus as research topics trickle down into graduate-level teaching, first textbooks written for new, cutting-edge courses may make their way into *Universitext*.

For further volumes:
www.springer.com/series/223

Szymon Borak • Wolfgang Karl Härdle
Brenda López-Cabrera

Statistics of Financial Markets

Exercises and Solutions

Second Edition

Springer

Szymon Borak
Wolfgang Karl Härdle
Brenda López-Cabrera
Humboldt-Universität zu Berlin
Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. Centre for Applied Statistics and Economics
School of Business and Economics
Berlin
Germany

Quantlets may be downloaded from http://extras.springer.com or via a link on http://springer.com/978-3-642-33928-8 or www.quantlet.org for a repository of quantlets.

ISBN 978-3-642-33928-8 ISBN 978-3-642-33929-5 (eBook)
DOI 10.1007/978-3-642-33929-5
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012954542

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface to the Second Edition

More practice makes you even more perfect. Many readers of the first edition of this book have followed this advice. We have received very helpful comments of the users of our book and we have tried to make it more perfect by presenting you the second edition with more quantlets in Matlab and R and with more exercises, e.g., for Exotic Options (Chap. 9).

This new edition is a good complement for the third edition of *Statistics of Financial Markets*. It has created many financial engineering practitioners from the pool of students at C.A.S.E. at Humboldt-Universität zu Berlin. We would like to express our sincere thanks for the highly motivating comments and feedback on our quantlets. Very special thanks go to the *Statistics of Financial Markets* class of 2012 for their active collaboration with us. We would like to thank in particular Mengmeng Guo, Shih-Kang Chao, Elena Silyakova, Zografia Anastasiadou, Anna Ramisch, Matthias Fengler, Alexander Ristig, Andreas Golle, Jasmin Krauß, Awdesch Melzer, Gagandeep Singh and, last but not least, Derrick Kanngießer.

Berlin, Germany, January 2013

Szymon Borak
Wolfgang Karl Härdle
Brenda López Cabrera

Preface to the First Edition

Wir behalten von unseren Studien am Ende doch nur das, was wir praktisch anwenden.
"In the end, we really only retain from our studies that which we apply in a practical way."

J. W. Goethe, Gespräche mit Eckermann, 24. Feb. 1824.

The complexity of modern financial markets requires good comprehension of economic processes, which are understood through the formulation of statistical models. Nowadays one can hardly imagine the successful performance of financial products without the support of quantitative methodology. Risk management, option pricing and portfolio optimisation are typical examples of extensive usage of mathematical and statistical modelling. Models simplify complex reality; the simplification though might still demand a high level of mathematical fitness. One has to be familiar with the basic notions of probability theory, stochastic calculus and statistical techniques. In addition, data analysis, numerical and computational skills are a must.

Practice makes perfect. Therefore the best method of mastering models is working with them. In this book, we present a collection of exercises and solutions which can be helpful in the advanced comprehension of *Statistics of Financial Markets*. Our exercises are correlated to Franke, Härdle, and Hafner (2011). The exercises illustrate the theory by discussing practical examples in detail. We provide computational solutions for the majority of the problems. All numerical solutions are calculated with R and Matlab. The corresponding quantlets – a name we give to these program codes – are indicated by in the text of this book. They follow the name scheme SFSxyz123 and can be downloaded from the Springer homepage of this book or from the authors' homepages.

Financial markets are global. We have therefore added, below each chapter title, the corresponding translation in one of the world languages. We also head each section with a proverb in one of those world languages. We start with a German proverb from Goethe on the importance of practice.

We have tried to achieve a good balance between theoretical illustration and practical challenges. We have also kept the presentation relatively smooth and, for more detailed discussion, refer to more advanced text books that are cited in the reference sections.

The book is divided into three main parts where we discuss the issues relating to option pricing, time series analysis and advanced quantitative statistical techniques.

The main motivation for writing this book came from our students of the course *Statistics of Financial Markets* which we teach at the Humboldt-Universität zu Berlin. The students expressed a strong demand for solving additional problems and assured us that (in line with Goethe) giving plenty of examples improves learning speed and quality. We are grateful for their highly motivating comments, commitment and positive feedback. In particular we would like to thank Richard Song, Julius Mungo, Vinh Han Lien, Guo Xu, Vladimir Georgescu and Uwe Ziegenhagen for advice and solutions on LaTeX. We are grateful to our colleagues Ying Chen, Matthias Fengler and Michel Benko for their inspiring contributions to the preparation of lectures. We thank Niels Thomas from Springer-Verlag for continuous support and for valuable suggestions on the writing style and the content covered.

Berlin, Germany

Szymon Borak
Wolfgang Härdle
Brenda López Cabrera

Contents

Language List

Arabic	اللغة العربية
Chinese	中文
Colognian	**Kölsch**
Croatian	**Hrvatski jezik**
Czech	Čeština
Dutch	**Nederlands**
English	**English**
French	**Français**
German	**Deutsch**
Greek	ελληνική γλώσσα
Hebrew	עִבְרִית
Hindi	हिन्दी
Indonesian	**Indonesia**
Italian	**Italiano**
Japanese	日本語

Korean	한국말
Latin	**lingua Latina**
Polish	język polski
Romanian	**Român**
Russian	русский язык
Spanish	**Español**
Ukrainian	украї́нська
Vietnamese	tiếng Việt

Symbols and Notation

Basics

X, Y	random variables or vectors
$X_1, X_2, \ldots, X_p$	random variables
$X = (X_1, \ldots, X_p)^\top$	random vector
$X \sim \cdot$	X has distribution $\cdot$
Γ, Δ	matrices
Σ	covariance matrix
1_n	vector of ones $(\underbrace{1, \ldots, 1}_{n\text{-times}})^\top$
0_n	vector of zeros $(\underbrace{0, \ldots, 0}_{n\text{-times}})^\top$
$\mathcal{I}_p$	identity matrix
$\mathbf{1}(.)$	indicator function, for a set M is $\mathbf{1} = 1$ on M, $\mathbf{1} = 0$ otherwise
$\mathbf{i}$	$\sqrt{-1}$
$\Rightarrow$	implication
$\Leftrightarrow$	equivalence
$\approx$	approximately equal
$\otimes$	Kronecker product
iff	if and only if, equivalence
SDE	stochastic differential equation
W_t	standard Wiener process
$\mathbb{N}$	Positive integer set
$\mathbb{Z}$	Integer set
$(X)^+$	$\lvert X \rvert * \mathbf{1}(X > 0)$

$[\lambda]$	Largest integer not larger than λ
$a.s.$	almost surely
$\alpha_n = \mathcal{O}(\beta_n)$	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 $n \longrightarrow \infty$

Characteristics of Distribution

$f(x)$	pdf or density of X
$f(x, y)$	joint density of X and Y
$f_X(x), f_Y(y)$	marginal densities of X and Y
$f_{X_1}(x_1), \ldots, f_{X_p}(x_p)$	marginal densities of $X_1, \ldots, X_p$
$\hat{f}_h(x)$	histogram or kernel estimator of $f(x)$
$F(x)$	cdf or distribution function of X
$F(x, y)$	joint distribution function of X and Y
$F_X(x), F_Y(y)$	marginal distribution functions of X and Y
$F_{X_1}(x_1), \ldots, F_{X_p}(x_p)$	marginal distribution functions of $X_1, \ldots, X_p$
$f_{Y\|X=x}(y)$	conditional density of Y given $X = x$
$\varphi_X(t)$	characteristic function of X
m_k	kth moment of X
κ_j	cumulants or semi-invariants of X

Moments

$\mathsf{E}\,X, \mathsf{E}\,Y$	mean values of random variables or vectors X and Y
$\mathsf{E}(Y\|X = x)$	conditional expectation of random variable or vector Y given $X = x$
$\mu_{Y\|X}$	conditional expectation of Y given X
$\mathsf{Var}(Y\|X = x)$	conditional variance of Y given $X = x$
$\sigma^2_{Y\|X}$	conditional variance of Y given X
$\sigma_{XY} = \mathsf{Cov}(X, Y)$	covariance between random variables X and Y
$\sigma_{XX} = \mathsf{Var}(X)$	variance of random variable X
$\rho_{XY} = \dfrac{\mathsf{Cov}(X, Y)}{\sqrt{\mathsf{Var}(X)\,\mathsf{Var}(Y)}}$	correlation between random variables X and Y
$\Sigma_{XY} = \mathsf{Cov}(X, Y)$	covariance between random vectors X and Y, i.e., $\mathsf{Cov}(X, Y) = E(X - EX)(Y - EY)^\top$
$\Sigma_{XX} = \mathsf{Var}(X)$	covariance matrix of the random vector X

Samples

x, y	observations of X and Y
$x_1, \ldots, x_n = \{x_i\}_{i=1}^n$	sample of n observations of X
$\mathcal{X} = \{x_{ij}\}_{i=1,\ldots,n; j=1,\ldots,p}$	$(n \times p)$ data matrix of observations of $X_1, \ldots, X_p$ or of $X = (X_1, \ldots, X_p)^\top$
$x_{(1)}, \ldots, x_{(n)}$	the order statistic of $x_1, \ldots, x_n$

Empirical Moments

$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	average of X sampled by $\{x_i\}_{i=1,\ldots,n}$
$s_{XY} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$	empirical covariance of random variables X and Y sampled by $\{x_i\}_{i=1,\ldots,n}$ and $\{y_i\}_{i=1,\ldots,n}$
$s_{XX} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$	empirical variance of random variable X sampled by $\{x_i\}_{i=1,\ldots,n}$
$r_{XY} = \frac{s_{XY}}{\sqrt{s_{XX}s_{YY}}}$	empirical correlation of X and Y
$\mathcal{S} = \{s_{X_iX_j}\}$	empirical covariance matrix of $X_1, \ldots, X_p$ or of the random vector $X = (X_1, \ldots, X_p)^\top$
$\mathcal{R} = \{r_{X_iX_j}\}$	empirical correlation matrix of $X_1, \ldots, X_p$ or of the random vector $X = (X_1, \ldots, X_p)^\top$

Distributions

$\varphi(x)$	density of the standard normal distribution
$\Phi(x)$	distribution function of the standard normal distribution
$\mathrm{N}(0, 1)$	standard normal or Gaussian distribution
$\mathrm{N}(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$\mathrm{N}_p(\mu, \Sigma)$	p-dimensional normal distribution with mean μ and covariance matrix Σ
$B(n, p)$	binomial distribution with parameters n and p
$lognormal(\mu, \sigma^2)$	lognormal distribution with mean μ and variance σ^2
$\xrightarrow{\mathcal{L}}$	convergence in distribution

$\xrightarrow{P}$	convergence in probability
CLT	Central Limit Theorem
χ^2_p	χ^2 distribution with p degrees of freedom
$\chi^2_{1-\alpha;p}$	$1-\alpha$ quantile of the χ^2 distribution with p degrees of freedom
t_n	t-distribution with n degrees of freedom
$t_{1-\alpha/2;n}$	$1-\alpha/2$ quantile of the t-distribution with n degrees of freedom
$F_{n,m}$	F-distribution with n and m degrees of freedom
$F_{1-\alpha;n,m}$	$1-\alpha$ quantile of the F-distribution with n and m degrees of freedom

Mathematical Abbreviations

$\mathrm{tr}(\mathcal{A})$	trace of matrix $\mathcal{A}$
$\mathrm{diag}(\mathcal{A})$	diagonal of matrix $\mathcal{A}$
$\mathrm{rank}(\mathcal{A})$	rank of matrix $\mathcal{A}$
$\det(\mathcal{A})$ or $\lvert\mathcal{A}\rvert$	determinant of matrix $\mathcal{A}$
$\mathrm{hull}(x_1,\ldots,x_k)$	convex hull of points $\{x_1,\ldots,x_k\}$
$\mathrm{span}(x_1,\ldots,x_k)$	linear space spanned by $\{x_1,\ldots,x_k\}$

Financial Market Terminology

OTC	over-the-counter
$self-financing$	a portfolio strategy with no resulting cash flow
$riskmeasure$	a mapping from a set of random variables (representing the risk at hand) to the real numbers

Some Terminology

Кто не рискует, тот не пьёт шампанского.
No pains, no gains.

This section contains an overview of some terminology that is used throughout the book. The notations are in part identical to those of Harville (2001). More detailed definitions and further explanations of the statistical terms can be found, e.g., in Breiman (1973), Feller (1966), Härdle and Simar (2012), Mardia, Kent, and Bibby (1979), or Serfling (2002).

adjoint matrix The *adjoint matrix* of an $n \times n$ matrix $\mathcal{A} = \{a_{ij}\}$ is the transpose of the cofactor matrix of $\mathcal{A}$ (or equivalently is the $n \times n$ matrix whose ijth element is the cofactor of a_{ji}).

asymptotic normality A sequence $X_1, X_2, \ldots$ of random variables is *asymptotically normal* if there exist sequences of constants $\{\mu_i\}_{i=1}^{\infty}$ and $\{\sigma_i\}_{i=1}^{\infty}$ such that $\sigma_n^{-1}(X_n - \mu_n) \xrightarrow{\mathcal{L}} N(0,1)$. The asymptotic normality means that for sufficiently large n, the random variable X_n has approximately $\mathrm{N}(\mu_n, \sigma_n^2)$ distribution.

bias Consider a random variable X that is parametrized by $\theta \in \Theta$. Suppose that there is an estimator $\widehat{\theta}$ of θ. The *bias* is defined as the systematic difference between $\widehat{\theta}$ and θ, $\mathsf{E}\{\widehat{\theta} - \theta\}$. The estimator is unbiased if $\mathsf{E}\,\widehat{\theta} = \theta$.

characteristic function Consider a random vector $X \in \mathbb{R}^p$ with pdf f. The *characteristic function* (cf) is defined for $t \in \mathbb{R}^p$:

$$\varphi_X(t) - \mathsf{E}[\exp(it^\top X)] = \int \exp(it^\top X) f(x) dx.$$

The cf fulfills $\varphi_X(0) = 1$, $|\varphi_X(t)| \leq 1$. The pdf (density) f may be recovered from the cf: $f(x) = (2\pi)^{-p} \int \exp(-it^\top X)\varphi_X(t)dt$.

characteristic polynomial (and equation) Corresponding to any $n \times n$ matrix $\mathcal{A}$ is its characteristic polynomial, say $p(.)$, defined (for $-\infty < \lambda < \infty$) by $p(\lambda) = |\mathcal{A} - \lambda\mathcal{I}|$, and its characteristic equation $p(\lambda) = 0$ obtained by setting its characteristic polynomial equal to 0; $p(\lambda)$ is a polynomial in λ of degree n and hence is of the form $p(\lambda) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + c_n\lambda^n$, where the coefficients $c_0, c_1, \ldots, c_{n-1}, c_n$ depend on the elements of $\mathcal{A}$.

conditional distribution Consider the joint distribution of two random vectors $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ with pdf $f(x, y) : \mathbb{R}^{p+1} \longrightarrow \mathbb{R}$. The marginal density of X is $f_X(x) = \int f(x, y)dy$ and similarly $f_Y(y) = \int f(x, y)dx$. The *conditional density* of X given Y is $f_{X|Y}(x|y) = f(x, y)/f_Y(y)$. Similarly, the conditional density of Y given X is $f_{Y|X}(y|x) = f(x, y)/f_X(x)$.

conditional moments Consider two random vectors $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ with joint pdf $f(x, y)$. The *conditional moments* of Y given X are defined as the moments of the conditional distribution.

contingency table Suppose that two random variables X and Y are observed on discrete values. The two-entry frequency table that reports the simultaneous occurrence of X and Y is called a *contingency table*.

critical value Suppose one needs to test a hypothesis $H_0 : \theta = \theta_0$. Consider a test statistic T for which the distribution under the null hypothesis is given by P_{θ_0}. For a given significance level α, the *critical value* is c_α such that $\mathrm{P}_{\theta_0}(T > c_\alpha) = \alpha$. The critical value corresponds to the threshold that a test statistic has to exceed in order to reject the null hypothesis.

cumulative distribution function (cdf) Let X be a p-dimensional random vector. The *cumulative distribution function* (cdf) of X is defined by $F(x) = \mathrm{P}(X \le x) = \mathrm{P}(X_1 \le x_1, X_2 \le x_2, \ldots, X_p \le x_p)$.

eigenvalues and eigenvectors An *eigenvalue* of an $n \times n$ matrix $\mathcal{A}$ is (by definition) a scalar (real number), say λ, for which there exists an $n \times 1$ vector, say x, such that $\mathcal{A}x = \lambda x$, or equivalently such that $(\mathcal{A} - \lambda\mathcal{I})x = \mathbf{0}$; any such vector x is referred to as an *eigenvector* (of $\mathcal{A}$) and is said to belong to (or correspond to) the eigenvalue λ. Eigenvalues (and eigenvectors), as defined herein, are restricted to real numbers (and vectors of real numbers).

eigenvalues (not necessarily distinct) The characteristic polynomial, say $p(.)$, of an $n \times n$ matrix $\mathcal{A}$ is expressible as

$$p(\lambda) = (-1)^n(\lambda - d_1)(\lambda - d_2)\cdots(\lambda - d_m)q(\lambda) \qquad (-\infty < \lambda < \infty),$$

where $d_1, d_2, \ldots, d_m$ are not-necessarily-distinct scalars and $q(.)$ is a polynomial (of degree $n - m$) that has no real roots; $d_1, d_2, \ldots, d_m$ are referred to as the *not-necessarily-distinct eigenvalues* of $\mathcal{A}$ or (at the possible risk of confusion) simply as the eigenvalues of $\mathcal{A}$. If the spectrum of $\mathcal{A}$ has k members, say $\lambda_1, \ldots, \lambda_k$, with algebraic multiplicities of $\gamma_1, \ldots, \gamma_k$, respectively, then $m = \sum_{i=1}^k \gamma_i$, and (for $i = 1, \ldots, k$) γ_i of the m not-necessarily-distinct eigenvalues equal λ_i .

empirical distribution function Assume that $X_1, \ldots, X_n$ are iid observations of a p-dimensional random vector. The *empirical distribution function* (edf) is defined through $F_n(x) = n^{-1} \sum_{i=1}^n \mathbf{1}(X_i \le x)$.

empirical moments The moments of a random vector X are defined through $m_k = \mathsf{E}(X^k) = \int x^k dF(x) = \int x^k f(x)dx$. Similarly, the *empirical moments* are defined through the empirical distribution function $F_n(x) = n^{-1} \sum_{i=1}^{n} \mathbf{1}(X_i \leq x)$. This leads to $\widehat{m}_k = n^{-1} \sum_{i=1}^{n} X_i^k = \int x^k dF_n(x)$.

estimate An *estimate* is a function of the observations designed to approximate an unknown parameter value.

estimator An *estimator* is the prescription (on the basis of a random sample) of how to approximate an unknown parameter.

expected (or mean) value For a random vector X with pdf f the *mean* or *expected value* is $\mathsf{E}(X) = \int x f(x)dx$.

Hessian matrix The *Hessian matrix* of a function f, with domain in $\mathbb{R}^{m\times 1}$, is the $m \times m$ matrix whose ijth element is the ijth partial derivative $D^2_{ij} f$ of f.

kernel density estimator The *kernel density estimator* $\widehat{f}_h$ of a pdf f, based on a random sample $X_1, X_2, \ldots, X_n$ from f, is defined by

$$\widehat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right).$$

The properties of the estimator $\widehat{f}_h(x)$ depend on the choice of the kernel function $K(.)$ and the bandwidth h. The kernel density estimator can be seen as a smoothed histogram; see also Härdle, Müller, Sperlich, and Werwatz (2004).

likelihood function Suppose that $\{x_i\}_{i=1}^n$ is an iid sample from a population with pdf $f(x;\theta)$. The *likelihood function* is defined as the joint pdf of the observations $x_1, \ldots, x_n$ considered as a function of the parameter θ, i.e., $L(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i;\theta)$. The log-likelihood function, $\ell(x_1, \ldots, x_n;\theta) = \log L(x_1, \ldots, x_n;\theta) = \sum_{i=1}^{n} \log f(x_i;\theta)$, is often easier to handle.

linear dependence or independence A nonempty (but finite) set of matrices (of the same dimensions $(n \times p)$), say $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_k$, is (by definition) *linearly dependent* if there exist scalars $x_1, x_2, \ldots, x_k$, not all 0, such that $\sum_{i=1}^{k} x_i \mathcal{A}_i = 0_n 0_p^\top$; otherwise (if no such scalars exist), the set is linearly independent. By convention, the empty set is linearly independent.

marginal distribution For two random vectors X and Y with the joint pdf $f(x, y)$, the *marginal pdfs* are defined as $f_X(x) = \int f(x, y)dy$ and $f_Y(y) = \int f(x, y)dx$.

marginal moments The *marginal moments* are the moments of the marginal distribution.

mean The *mean* is the first-order empirical moment $\overline{x} = \int x dF_n(x) = n^{-1} \sum_{i=1}^{n} x_i = \widehat{m}_1$.

mean squared error (MSE) Suppose that for a random vector C with a distribution parametrized by $\theta \in \Theta$ there exists an estimator $\widehat{\theta}$. The *mean squared error* (MSE) is defined as $\mathsf{E}_X(\widehat{\theta} - \theta)^2$.

median Suppose that X is a continuous random variable with pdf $f(x)$. The *median* $\widetilde{x}$ lies in the center of the distribution. It is defined as $\int_{-\infty}^{\widetilde{x}} f(x)dx = \int_{\widetilde{x}}^{+\infty} f(x)dx = 0.5$.

moments The *moments* of a random vector X with the distribution function $F(x)$ are defined through $m_k = \mathsf{E}(X^k) = \int x^k dF(x)$. For continuous random vectors with pdf $f(x)$, we have $m_k = \mathsf{E}(X^k) = \int x^k f(x)dx$.

normal (or Gaussian) distribution A random vector X with the *multinormal distribution* $N(\mu, \Sigma)$ with the mean vector μ and the variance matrix Σ is given by the pdf

$$f_X(x) = |2\pi\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right\}.$$

orthogonal matrix An $(n \times n)$ matrix $\mathcal{A}$ is *orthogonal* if $\mathcal{A}^\top\mathcal{A} = \mathcal{A}\mathcal{A}^\top = \mathcal{I}_n$.

probability density function (pdf) For a continuous random vector X with cdf F, the *probability density function* (pdf) is defined as $f(x) = \partial F(x)/\partial x$.

quantile For a random variable X with pdf f the α *quantile* q_α is defined through: $\int_{-\infty}^{q_\alpha} f(x)dx = \alpha$.

p**-value** The critical value c_α gives the critical threshold of a test statistic T for rejection of a null hypothesis $H_0 : \theta = \theta_0$. The probability $\mathrm{P}_{\theta_0}(T > c_\alpha) = p$ defines that *p-value*. If the p-value is smaller than the significance level α, the null hypothesis is rejected.

random variable and vector Random events occur in a probability space with a certain even structure. A *random variable* is a function from this probability space to $\mathbb{R}$ (or $\mathbb{R}^p$ for random vectors) also known as the state space. The concept of a random variable (vector) allows one to elegantly describe events that are happening in an abstract space.

scatterplot A *scatterplot* is a graphical presentation of the joint empirical distribution of two random variables.

singular value decomposition (SVD) An $m \times n$ matrix $\mathcal{A}$ of rank r is expressible as

$$\mathcal{A} = \mathcal{P}\begin{pmatrix} \mathcal{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\mathcal{Q}^\top = \mathcal{P}_1\mathcal{D}_1\mathcal{Q}_1^\top = \sum_{i=1}^{r} s_i p_i q_i^\top = \sum_{j=1}^{k} \alpha_j \mathcal{U}_j\,,$$

where $\mathcal{Q} = (q_1, \ldots, q_n)$ is an $n \times n$ orthogonal matrix and $\mathcal{D}_1 = \mathrm{diag}(s_1, \ldots, s_r)$ an $r \times r$ diagonal matrix such that $\mathcal{Q}^\top\mathcal{A}^\top\mathcal{A}\mathcal{Q} = \begin{pmatrix} \mathcal{D}_1^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, where $s_1, \ldots, s_r$ are (strictly) positive, where $\mathcal{Q}_1 = (q_1, \ldots, q_r)$, $\mathcal{P}_1 = (p_1, \ldots, p_r) = \mathcal{A}\mathcal{Q}_1\mathcal{D}_1^{-1}$, and, for any $m \times (m - r)$ matrix $\mathcal{P}_2$ such that $\mathcal{P}_1^\top\mathcal{P}_2 = \mathbf{0}$, $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$, where $\alpha_1, \ldots, \alpha_k$ are the distinct values represented among $s_1, \ldots, s_r$, and where (for $j = 1, \ldots, k$) $\mathcal{U}_j = \sum_{\{i\,:\,s_i=\alpha_j\}} p_i q_i^\top$; any of these four representations may be referred to as the *singular value decomposition* of $\mathcal{A}$, and $s_1, \ldots, s_r$ are referred to as the singular values of $\mathcal{A}$. In fact, $s_1, \ldots, s_r$ are the positive square roots of the nonzero eigenvalues of $\mathcal{A}^\top\mathcal{A}$ (or equivalently $\mathcal{A}\mathcal{A}^\top$), $q_1, \ldots, q_n$ are eigenvectors of $\mathcal{A}^\top\mathcal{A}$, and the columns of $\mathcal{P}$ are eigenvectors of $\mathcal{A}\mathcal{A}^\top$.

spectral decomposition A $p \times p$ symmetric matrix $\mathcal{A}$ is expressible as

$$\mathcal{A} = \Gamma \Lambda \Gamma^\top = \sum_{i=1}^{p} \lambda_i \gamma_i \gamma_i^\top$$

where $\lambda_1, \ldots, \lambda_p$ are the not-necessarily-distinct eigenvalues of $\mathcal{A}$, $\gamma_1, \ldots, \gamma_p$ are orthonormal eigenvectors corresponding to $\lambda_1, \ldots, \lambda_p$, respectively, $\Gamma = (\gamma_1, \ldots, \gamma_p)$, $\mathcal{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$.

subspace A *subspace* of a linear space $\mathcal{V}$ is a subset of $\mathcal{V}$ that is itself a linear space.

Taylor expansion The *Taylor series* of a function $f(x)$ in a point a is the power series $\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n$. A truncated Taylor series is often used to approximate the function $f(x)$.

List of Figures

Part I
Option Pricing

Chapter 1
Derivatives

파생상품

달걀을 한 바구니에 담지 마라.
Don't put all eggs in one basket

A derivative (derivative security or contingent claim) is a financial instrument whose value depends on the value of others, more basic underlying variables. Options, future contracts, forward contracts, and swaps are examples of derivatives. The aim of this chapter is to present and discuss various options strategies. The exercises emphasize the differences of the strategies through an intuitive approach using payoff graphs.

Exercise 1.1 (Butterfly strategy). *Consider a butterfly strategy: a long call option with an exercise price of 100 USD, a second long call option with an exercise price of 120 USD and two short calls with an exercise price of 110 USD. Give the payoff table for different stock values. When will this strategy be preferred?*

The payoff table for different stock values:

Strategy	$S_T \leq 100$	$100 < S_T \leq 110$	$110 < S_T \leq 120$	$120 < S_T$
A long call at 100	0	$S_T - 100$	$S_T - 100$	$S_T - 100$
A long call at 120	0	0	0	$S_T - 120$
Two short calls at 110	0	0	$2(110 - S_T)$	$2(110 - S_T)$
Total	0	$S_T - 100$	$120 - S_T$	0

This strategy is preferred when the stock price fluctuates slightly around 110 USD.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_1,

Exercise 1.2 (Risk of a strategy). *Consider a simple strategy: an investor buys one stock, one European put with an exercise price K, sells one European call with an exercise price K. Calculate the payoff and explain the risk of this strategy.*

Strategy	$S_T \leq K$	$S_T > K$
Buy a stock	S_T	S_T
Buy a put	$K - S_T$	0
Sell a call	0	$-(S_T - K)$
Total	K	K

This is a risk-free strategy. The value of portfolio at time T is the exercise price K, which is not dependent on the stock price at expiration date.

Exercise 1.3 (Bull call spread). *One of the most popular types of the spreads is a bull spread. A bull-call-price spread can be made by buying a call option with a certain exercise price and selling a call option on the same stock with a higher exercise price. Both call options have the same expiration date. Consider a European call with an exercise price of K_1 and a second European call with an exercise price of K_2. Draw the payoff table and payoff graph for this strategy.*

Strategy	$S_T \leq K_1$	$K_1 < S_T \leq K_2$	$K_2 < S_T$
A long call at K_1	0	$S_T - K_1$	$S_T - K_1$
A short call at K_2	0	0	$K_2 - S_T$
Total	0	$S_T - K_1$	$K_2 - K_1$

Suppose that a trader buys a call for 12 USD with a strike price of $K_1 = 100$ USD and sells a call for 8 USD with a strike price of $K_2 = 120$ USD. If the stock price is above 120 USD, the payoff from this strategy is 16 USD (8 USD from short call, 8 USD from long call). The cost of this strategy is 4 USD (buying a call for 12 USD, selling a call for 8 USD). If the stock price is between 100 and 120 USD, the payoff is $S_T - 104$. The bull spread strategy limits the trader's upside as well as downside risk. The payoff graph for the bull call spread strategy is shown in Fig. 1.1.

Exercise 1.4 (Straddle). *Consider a strategy of buying a call and a put with the same strike price and expiration date. This strategy is called straddle. The price of the long call option is 3 USD. The price of the long put option is 5 USD. The strike price is $K = 40$ USD. Draw the payoff table and payoff graph for the straddle strategy (Fig. 1.2).*

The advantage of a straddle is that the investor can profit from stock prices moving in both directions. One does not care whether the stock price goes up or down, but only how much it moves. The disadvantage to a straddle is that it has a high premium because of having to buy two options. The initial cost of the straddle at a stock price 40 USD is 8 USD (3 USD for the call and 5 USD for the put). If

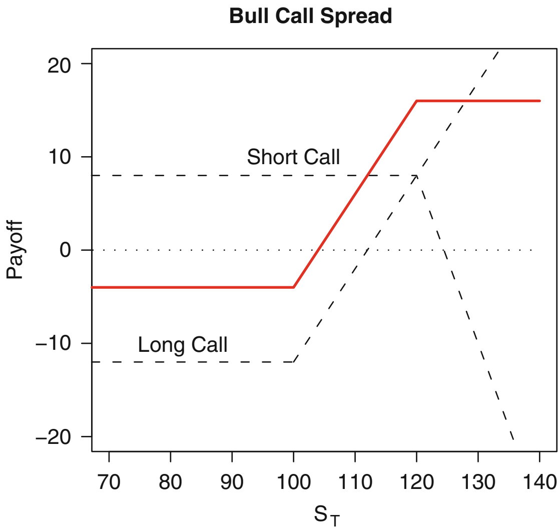


Fig. 1.1 Bull call spread SFSbullspreadcall

Payoff	$S_T \leq K$	$S_T > K$
Payoff from call	0	$S_T - K$
Payoff from put	$K - S_T$	0
Total payoff	$K - S_T$	$S_T - K$

the stock price stays at 38 USD, we can see that the strategy costs the trader 6 USD. Since the initial cost is 8 USD, the call expires worthless, and the put expires worth 2 USD. However, if the stock price jumps to 60 USD, a profit of 12 USD (60-40-8) is made. If the stock price goes down to 30 USD, a profit of 2 USD (40-30-8) is made, and so on. The payoff graph for the straddle option strategy is shown in Fig. 1.3.

Exercise 1.5 (Butterfly spread). *Consider the option spread strategy known as the butterfly spread. A butterfly spread involves positions in options with three different strike prices. It can be created by buying a call option with a relatively low strike price K_1, buying a call option with a relatively high strike price K_3, and selling two call options with a strike price $K_2 = 0.5(K_1 + K_3)$. Draw the payoff table and payoff graph for the butterfly spread strategy.*

Position	$S_T \leq K_1$	$K_1 < S_T \leq K_2$	$K_2 < S_T \leq K_3$	$S_T > K_3$
First long call	0	$S_T - K_1$	$S_T - K_1$	$S_T - K_1$
Second long call	0	0	0	$S_T - K_3$
Two short calls	0	0	$-2(S_T - K_2)$	$-2(S_T - K_2)$
Total payoff	0	$S_T - K_1$	$K_3 - S_T$	0

Fig. 1.2 Example of a straddle with the S&P 500 index as underlying

Suppose that the market prices of 3-month calls are as follows:

Strike price (USD)	Price of call (USD)
65	12
70	8
75	5

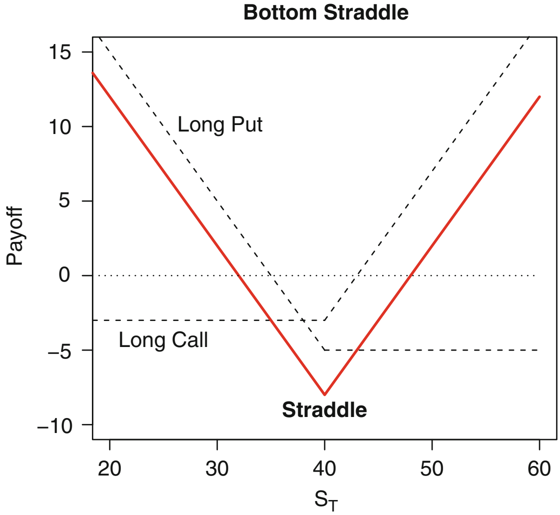


Fig. 1.3 Bottom straddle `SFSbottomstraddle`

A trader could create a butterfly spread by buying one call with a strike price of 65 USD, buying one call with a strike price of 75 USD, and selling two calls with a strike price of 70 USD. It costs $12 + 5 - 2 * 8 = 1$ USD to create this spread. If the stock price in 3 months is greater than 75 USD or less than 65 USD, the trader will lose 1 USD. If the stock price is between 66 and 74 USD, the trader will make a profit. The maximum profit is reached if the stock price in 3 months is 70 USD. Hence, this strategy should be used if the trader thinks that the stock price will stay close to K_2 in the future. The payoff graph for the butterfly spread using call options is shown in Fig. 1.4.

Exercise 1.6 (Butterfly spread). *Butterfly spreads can be implemented using put options. If put contracts are used, the strategy would necessitate two long put contracts, one with a low strike price* K_1 *and a second with a higher strike price* K_3*, and two short puts with a strike price* $K_2 = 0.5(K_1 + K_3)$. *Draw payoff graph for the butterfly spread using put options.*

Suppose that the market prices of 3-month puts are as follows:

Strike price (USD)	Price of put (USD)
65	5
70	8
75	12

The payoff graph for the butterfly spread using put options is shown in Fig. 1.5.

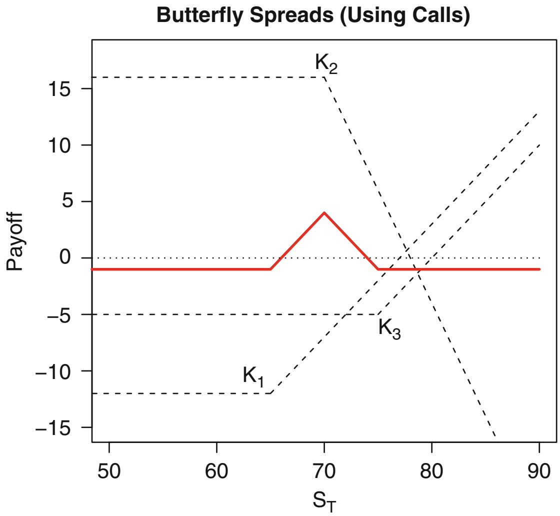


Fig. 1.4 Butterfly spread created using call options SFSbutterfly

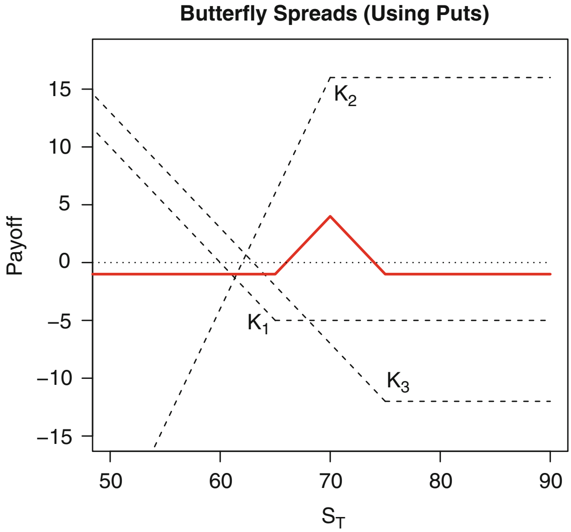


Fig. 1.5 Butterfly spread created using put options SFSbutterfly

Exercise 1.7 (Strangle). *Consider the option combination strategy known as the strangle. In the strangle strategy a trader buys a put and a call with a different strike price and the same expiration date. The put strike price,* K_1 *is smaller than the call strike price,* K_2*. Draw the payoff table and payoff graph for the strangle strategy.*

Position	$S_T \leq K_1$	$K_1 < S_T < K_2$	$K_2 \leq S_T$
Profit from call	0	0	$S_T - K_2$
Profit from put	$K_1 - S_T$	0	0
Total profit	$K_1 - S_T$	0	$S_T - K_2$

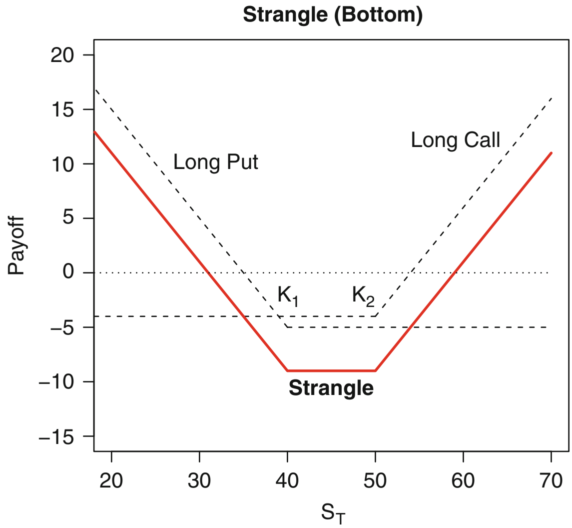


Fig. 1.6 Bottom strangle `SFSbottomstrangle`

The aim of the strangle strategy is to profit from an anticipated upward or downward movement in the stock price. The trader thinks there will be a large price movement but is not sure whether it will be an increase or decrease in price. The risk is minimized at a level between K_1 and K_2. Suppose that the put price is 5 USD with a strike price $K_1 = 40$ USD, the call price is 4 USD with a strike price $K_2 = 50$ USD. The payoff graph for the strangle strategy is shown in Fig. 1.6.

Exercise 1.8 (Strip). *Consider the option combination strategy known as a strip. A strip consists of one long call and two long puts with the same strike price and expiration date. Draw the payoff diagram for this option strategy.*

The aim of the strip is to profit from a large anticipated decline in the stock price below the strike price. Consider a strip strategy in which two long puts with the price of 3 USD for each and a long call with the price of 4 USD are purchased simultaneously with strike price $K = 35$ USD. The payoff graph for the strip strategy is shown in Fig. 1.7.

Exercise 1.9 (Strap). *Consider the option strategy known as a strap. A strap could be intuitively interpreted as the reverse of a strip. A strap consists of two long calls and one long put with same strike price and expiration date. Draw the payoff diagram for this option strategy.*

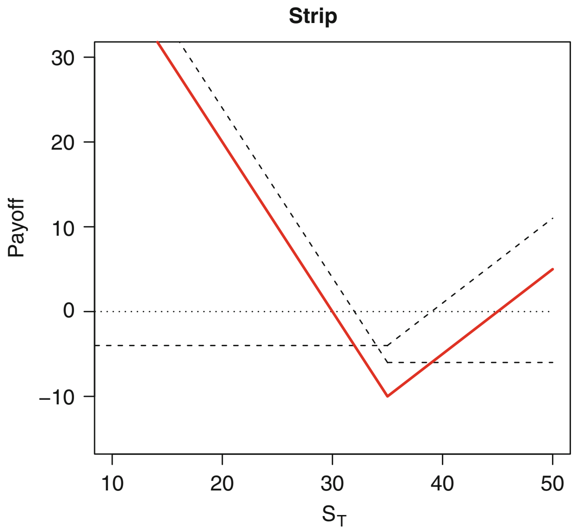


Fig. 1.7 Strip SFSstrip

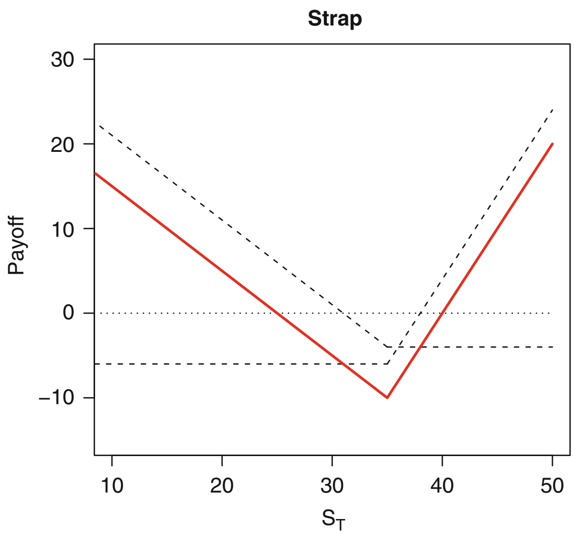


Fig. 1.8 Strap SFSstrap

The aim of the strap is to profit from a large anticipated rise in the stock price above strike price. The following payoff graph is drawn with two long call options, $C_0 = 3$ USD and one long put option, $P_0 = 4$ USD. The strike price is $K = 35$ USD for both options. The payoff graph for strap strategy is shown in Fig. 1.8.

Fig. 1.9 S&P 500 index for 2008

Exercise 1.10 (Choosing a Strategy). *The Bloomberg screenshot depicting the S&P 500 index in Fig. 1.9 illustrates the rapid decline in stock prices in the fall of 2008. Name possible strategies to make profit from such a downturn. What is decisive for choosing a strategy?*

Under circumstances like in the fall of 2008, several strategies can be thought of to make profit. Among those strategies are bull call spread, bear spread created using put options, bottom straddle, butterfly spread created using call options and butterfly spread created using puts options. The expectation formation about the future price developments determines which strategy should be chosen.

Exercise 1.11 (Straddle). *You are long a straddle with strike price $K = 25$ USD and price $S_t = 25$. The straddle costs you 5 USD to enter. What price movements are you looking for in the underlying?*

A straddle is a long call plus long put with the same strike price. If you hold the straddle until maturity, then you need a price change of more than 5 USD either way in the underlying in order to profit. A smaller price change, however, can lead to profits if it occurs before maturity.

Exercise 1.12 (Butterfly spread). *Call options on a stock are available with strike prices $K_1 = 15$ USD, $K_2 = 17.5$ USD, $K_3 = 20$ USD and time to maturity in 3 months. The prices are 4, 2 and 0.5 USD respectively. Explain how the options* *can be used to create a butterfly spread. Construct a payoff table that shows how profit varies with stock prices for the butterfly spread.*

A butterfly spread can be created by buying call options with strike prices $K_1 = 15$ USD and $K_3 = 20$ USD and by shorting two call options with strike prices $K_2 = 17.5$ USD. The total investment is $4 + 0.5$ USD $- 2 \cdot 2$ USD $= 0.5$ USD.

Position	$S_T \leq 15$	$15 < S_T \leq 17.5$	$17.5 < S_T \leq 20$	$S_T > 20$
First long call	-4	$(S_T - 15) - 4$	$(S_T - 15) - 4$	$(S_T - 15) - 4$
Second long call	-0.5	-0.5	-0.5	$(S_T - 20) - 0.5$
Two short calls	4	4	$-2(S_T - 17.5) + 4$	$-2(S_T - 17.5) + 4$
Total payoff	-0.5	$(S_T - 15) - 0.5$	$(20 - S_T) - 0.5$	-0.5

Chapter 2
Introduction to Option Management

Εισαγωγή στη Διαχείριση Επιλογή

χρὴ δ' ἐπ' ἀξίοις πονεῖν ψυχὴν προβάλλοντ' ἐν κύβοισι δαίμονος
The prize must be worth the toil when one stakes one's life on fortune's dice.

Dolon to Hector, Euripides (Rhesus, 182)

In this chapter we discuss basic concepts of option management. We will consider both European and American call and put options and practice concepts of pricing, look at arbitrage opportunities and the valuation of forward contracts. Finally, we will investigate the put-call parity relation for several cases.

Exercise 2.1 (Call and Put Options). *A company's stock price is $S_0 = 110$ USD today. It will either rise or fall by 20 % after one period. The risk-free interest rate for one period is $r = 10$ %.*

(a) Find the risk-neutral probability that makes the expected return of the asset equal to the risk-free rate.
(b) Find the prices of call and put options with the exercise price $K = 100$ USD.
(c) How can the put option be duplicated?
(d) How can the call option be duplicated?
(e) Check put-call parity.

(a) The risk-neutral probability in this one period binomial model satisfies

$$(1 + r)S_0 = \mathsf{E}_Q\, S_t,$$

where Q denotes the risk neutral (Bernoulli) measure with probability q. Plugging in the given data $S_0 = 110$, $S_{11} = 110 \cdot 1.2$, $S_{12} = 110 \cdot 0.8$ and $r = 0.1$ leads to:

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_2, © Springer-Verlag Berlin Heidelberg 2013

$$
\begin{aligned}
(1+0.1)\cdot 110 &= q\cdot 1.2\cdot 110 + (1-q)\cdot 0.8\cdot 110\\
1.1 &= 1.2q + 0.8(1-q)\\
0.3 &= 0.4q\\
q &= 0.75
\end{aligned}
$$

Hence the risk neutral probability measure Q is

$$
\begin{aligned}
P_Q(S_{11} &= 1.2\cdot 110) = 0.75\\
P_Q(S_{12} &= 0.8\cdot 110) = 0.25
\end{aligned}
$$

(b) The call option price is $C = (1+r)^{-1}\,\mathsf{E}_Q\,\Psi(S_1, K)$, with $K = 100$ and $\Psi(S,K) = 1(S-K>0)(S-K)$. Denote $c^u = \Psi(S_{11,K})$ and $c^d = \Psi(S_{12}, K)$. Then $C = \left\{qc^u + (1-q)c^d\right\}/(1+r)$ is the expected payoff discounted by the risk-free interest rate. Using the prior obtained values we know that the stock can either increase to $S_{11} = 110\cdot(1+0.2) = 132$ or decrease to $S_{12} = 110\cdot(1-0.2) = 88$, whereas the risk-neutral probability is $q = 0.75$. Given the exercise price of $K = 100$, the payoff in case of a stock price increase is $c^u = \max(132-100, 0) = 32$, in case of a decrease is $c^d = \max(88-100, 0) = 0$. Thus, the call price is $C = (0.75\cdot 32 + 0.25\cdot 0)/(1+0.1) = 21.82$ USD.

Then the put option price is calculated using $P = \left\{qp^u + (1-q)p^d\right\}/(1+r)$. Given the exercise price of $K = 100$ the payoff for a stock price increase is $p^u = \max(100-132, 0) = 0$ and for a decrease is $p^d = \max(100-88, 0) = 12$. Thus, the put price is $P = (0.75\cdot 0 + 0.25\cdot 12)/(1+0.1) = 2.73$ USD.

(c) Given an increase in the stock price, the value of the derivative is $p^u = \Delta S_{11} + \beta(1+r)$, where Δ is the the number of shares of the underlying asset, S_{11} is the value of the underlying asset at the top, β is the amount of money in the risk-free security and $1+r$ is the risk-free interest rate.

The value of p^d is calculated respectively as $p^d = \Delta S_{12} + \beta(1+r)$. Using $p^u = 0$, $p^d = 12$, $S_{11} = 132$ and $S_{12} = 88$ we can solve the two equations: $\Delta 132 + \beta(1+0.1) = 0$ and $\Delta 88 + \beta(1+0.1) = 12$ and obtain $\Delta = -0.27$, $\beta = 32.73$. This means that one should sell 0.27 shares of stock and invest 32.73 USD at the risk-free rate.

(d) For the call option, we can analogously solve the following two equations: $\Delta 132 + \beta(1+0.1) = 32$, $\Delta 88 + \beta(1+0.1) = 0$. Finally, we get $\Delta = 0.73$, $\beta = -58.18$. This means that one should buy 0.73 shares of stock and borrow 58.18 USD at a risk-free rate.

(e) The principle of put-call parity refers to the equivalence of the value of a European call and put option which have the same maturity date T, the same delivery price K and the same underlying. Hence, there are combinations of options which can create positions that are the same as holding the stock itself.

These option and stock positions must all have the same return or an arbitrage opportunity would be available to traders.

Formally, the relationship reads $C + K/(1+r) = P + S_0$. Refer to Franke et al. (2011) for the derivation. Plugging in the above calculated values yields $21.82 + 100/1.1 = 2.73 + 110$. Obviously, the equivalence holds, so the put-call parity is satisfied.

Exercise 2.2 (American Call Option). *Consider an American call option with a* 40 *USD strike price on a specific stock. Assume that the stock sells for* 45 *USD a share without dividends. The option sells for* 5 *USD 1 year before expiration. Describe an arbitrage opportunity, assuming the annual interest rate is 10 %.*

Short a share of the stock and use the 45 USD you receive to buy the option for 5 USD and place the remaining 40 USD in a savings account. The initial cash flow from this strategy is zero. If the stock is selling for more than 40 USD at expiration, exercise the option and use your savings account balance to pay the strike price. Although the stock acquisition is used to close out your short position, the $40 \cdot 0.1 = 4$ USD interest in the savings account is yours to keep. If the stock price is less than 40 USD at expiration, buy the stock with funds from the savings account to cancel the short position. The 4 USD interest in the savings account and the difference between the 40 USD (initial principal in the savings account) and the stock price is yours to keep (Table 2.1).

Exercise 2.3 (European Call Option). *Consider a* European call option *on a stock with current spot price* $S_0 = 20$*, dividend* $D = 2$ *USD, exercise price* $K = 18$ *and time to maturity* 6 *months. The annual risk-free rate is* $r = 10\,\%$*. What is the upper and lower bound (limit) of the price of the call and put options?*

The upper bound for a European call option is always the current market price of the stock S_0. If this is not the case, arbitrageurs could make a riskless profit by buying the stock and selling the call option. The upper limit for the call is therefore 20.

Based on $P + S_0 - K\exp(-r\tau) - D = C$ and $P \geq 0$, the lower bound for the price of a European call option is given by:

$$C \geq S_0 - K\exp(-r\tau) - D$$

$$C \geq 20 - 18\exp(-0.10 \cdot 6/12) - 2$$

$$C \geq 20 - 17.12 - 2$$

$$C \geq 0.88$$

Consider for example, a situation where the European call price is 0.5 USD. An arbitrageur could buy the call for 0.5 USD and short the stock for 20 USD. This provides a cash flow of $20 - 0.5 = 19.5$ USD which grows to $19.5\exp(0.1 \cdot 0.5) = 20.50$ in 6 months. If the stock price is greater than the exercise price at maturity,

Table 2.1 Cash flow table for this strategy

Action	CF_t	CF_T	
		$S_T < 40$	$S_T \geq 40$
Short a share of the stock	45	$-S_T$	$-S_T$
Buy a call	−5	0	$S_T - 40$
Rest to savings account	−40	44	44
Total	0	$(40 - S_T) + 4$	4

the arbitrageur will exercise the option, close out the short position and make a profit of $20.50 - 18 = 2.50$ USD.

If the price is less than 18 USD, the stock is bought in the market and the short position is closed out. For instance, if the price is 15 USD, the arbitrageur makes a profit of $20.50 - 15 = 5.50$ USD.

Thus, the price of the call option lies between 0.88 and 20 USD.

The upper bound for the put option is always the strike price $K = 18$ USD, while the lower bound is given by:

$$P \geq K \exp(-r\tau) - S_0 + D$$

$$P \geq 18 \exp(-0.10 \cdot 0.5) - 20 + 2$$

$$P \geq 17.12 - 20 + 2$$

$$P \geq -0.88$$

However, the put option price cannot be negative and therefore it can be further refined as:

$$P \geq \max\{K \exp(-r\tau) - S_0 + D, 0\}.$$

Thus, the price of this put option lies between 0 and 18 USD.

Exercise 2.4 (Spread between American Call and Put Option). *Assume that the above stock and option market data does not refer to European put and call options but rather to American put and call options. What conclusions can we draw about the relationship between the upper and lower bounds of the spread between the American call and put for a non-dividend paying stock?*

The relationship between the upper and lower bounds of the spread between American call and put options can be described by the following relationship: $S_0 - K \leq C - P \leq S_0 - K \exp(-r\tau)$. In this specific example, the spread between the prices of the American put and call options can be described as follows:

$$20 - 18 \leq C - P \leq 20 - 18 \exp(-0.10 \cdot 6/12)$$

$$2 \leq C - P \leq 2.88$$

Table 2.2 Portfolio value for some future time t'

	$S_{t'} \leq K_1$	$K_1 \leq S_{t'} \leq K_\lambda$	$K_\lambda \leq S_{t'} \leq K_0$	$K_0 \leq S_{t'}$
1.	$\lambda(K_1 - S_{t'})$	0	0	0
2.	$(1-\lambda)(K_0 - S_{t'})$	$(1-\lambda)(K_0 - S_{t'})$	$(1-\lambda)(K_0 - S_{t'})$	0
3.	$-(K_\lambda - S_{t'})$	$-(K_\lambda - S_{t'})$	0	0
Sum	0	$\lambda(S_{t'} - K_1)$	$(1-\lambda)(K_0 - S_{t'})$	0

Exercise 2.5 (Price of American and European Put Option). *Prove that the price of an American or European put option is a convex function of its exercise price.*

Additionally, consider two put options on the same underlying asset with the same maturity. The exercise prices and the prices of these two options are $K_1 = 80$ *and* 38.2 *EUR and* $K_2 = 50$ *and* 22.6 *EUR.*

There is a third put option on the same underlying asset with the same maturity. The exercise price of this option is 60 *EUR. What can be said about the price of this option?*

Let $\lambda \in [0, 1]$ and $K_1 < K_0$. Consider a portfolio with the following assets:

1. A long position in λ puts with exercise price K_1
2. A long position in $(1-\lambda)$ puts with exercise price K_0
3. A short position in 1 put with exercise price $K_\lambda \stackrel{\text{def}}{=} \lambda K_1 + (1-\lambda)K_0$

The value of this portfolio for some future time t' can be seen in Table 2.2:

The value of the portfolio is always bigger than or equal to 0. For no arbitrage to happen, the current value of the portfolio should also be non-negative, so:

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

The above inequality proves the convexity of the put option price with respect to its exercise price.

The price of a put option increases as the exercise price increases. So in this specific example:

$$P_{50,T} \leq P_{60,T} \leq P_{80,T}$$

and hence:

$$22.6 \leq P_{60,T} \leq 38.2$$

Moreover, we also know that the prices of call and put options are convex, so

$$\lambda K_1 + (1-\lambda)K_2 = 60$$
$$\lambda = 1/3$$

Table 2.3 Cash flow table for this strategy

Action	CF_t	CF_T	
		$S_T < 235$	$S_T \geq 235$
Buy a put	-5.25	$235 - S_T$	0
Short a call	21.88	0	$235 - S_T$
Buy a forward with $K = 235$	-16.17	$S_T - 235$	$S_T - 235$
Total	0.46	0	0

$$
\begin{aligned}
1/3 \cdot 38.2 + 2/3 \cdot 22.6 &\geq P_{60,T} \\
27.8 &\geq P_{60,T}
\end{aligned}
$$

The price of the put option in this example should be between 22.6 and 27.8 EUR

Exercise 2.6 (Put-Call Parity). *The present price of a stock without dividends is* 250 *EUR. The market value of a European call with strike price* 235 *EUR and time to maturity* 180 *days is* 21.88 *EUR. The annual risk-free rate is 1 %.*

(a) Assume that the market price for a European put with same strike price and time to maturity is 5.25 *EUR. Show that this is inconsistent with put-call parity.*
(b) Describe how you can take advantage of this situation by finding a combination of purchases and sales which provides an instant profit with no liability 180 days from now.

(a) Put-call parity gives:

$$
\begin{aligned}
P_{K,T}(S_t, \tau) &= C_{K,T}(S_t, \tau) - \{S_t - K \exp(-r\tau)\} \\
&= 21.88 - \{250 - 235 \exp(-0.01 \cdot 0.5)\} \\
&= 21.88 - 16.17 \\
&= 5.71
\end{aligned}
$$

Thus, the market value of the put is too low and it offers opportunities for arbitrage.

(b) Puts are underpriced, so we can make profit by buying them. We use CF_t to denote the cash flow at time t. The cash flow table for this strategy can be seen in Table 2.3.

Exercise 2.7 (Hedging Strategy). *A stock currently selling at S_0 with fixed dividend D_0 is close to its dividend payout date. Show that the parity value for the futures price on the stock can be written as $F_0 = S_0(1 + r)(1 - d)$, where $d = D_0/S_0$ and r is the risk-free interest rate for a period corresponding to the term of the futures contract. Construct an arbitrage table demonstrating the riskless strategy assuming that the dividend is reinvested in the stock. Is your result*

Table 2.4 Cash flow table for this strategy

Action	CF_0	CF_T
Buy one share immediately	$-S_0$	S_T
Reinvest the dividend	0	$S_T d/(1-d)$
Sell $1/(1-d)$ forwards	0	$(F_0 - S_T)/(1-d)$
Borrow S_0 euros	S_0	$-S_0(1+r)$
Total	0	$F_0/(1-d) - S_0(1+r)$

consistent with the parity value $F_0 = S_0(1+r) - FV(D_0)$ where the forward value $FV(x) = (1+r)x$? (Hint: How many shares will you hold after reinvesting the dividend? How will this affect your hedging strategy?)

The price of the stock will be $S_0(1-d)$ after the dividend has been paid, and the dividend amount will be dS_0. So the reinvested dividend could purchase $d/(1-d)$ shares of stock, and you end up with $1 + d/(1-d) = 1/(1-d)$ shares in total. You will need to sell that many forward contracts to hedge your position. Here is the strategy (Table 2.4):

To remove arbitrage, the final payoff should be zero, which implies:

$$\begin{aligned} F_0 &= S_0(1+r)(1-d) \\ &= S_0(1+r) - S_0(1+r)(D_0/S_0) \\ &= S_0(1+r) - D_0(1+r) \\ &= S_0(1+r) - FV(D_0) \end{aligned}$$

Exercise 2.8 (No-Arbitrage Theory). *Prove that the following relationship holds, using no-arbitrage theory.*

$$F(T_2) = F(T_1)(1+r)^{T_2 - T_1} - FV(D)$$

where $F_0(T)$ is today's futures price for delivery time T, $T_2 > T_1$, and FV(D) is the future value to which any dividends paid between $T_1$and T_2 will grow if invested risklessly until time T_2 (Table 2.5).

Since the cashflow at T_2 is riskless and no net investment is made, any profits would represent an arbitrage opportunity. Therefore, the zero-profit no-arbitrage restriction implies that

$$F(T_2) = F(T_1)(1+r)^{T_2 - T_1} - FV(D)$$

Exercise 2.9 (Arbitrage Opportunity). *Suppose that the current DAX index is 3,200, and the DAX index futures which matures exactly in 6 months are priced at 3,220.*

Table 2.5 Cash flow table for this strategy

Action	CF_0	CF_{T_1}	CF_{T_2}
Long futures with T_1 maturity	0	$S_1 - F(T_1)$	0
Short futures with T_2 maturity	0	0	$F(T_2) - S_2$
Buy the asset at T_1, sell at T_2.	0	$-S_1$	$S_2 + FV(D)$
Invest dividends paid until T_2			
At T_1, borrow $F(T_1)$	0	$F(T_1)$	$-F(T_1) \times (1+r)^{T_2-T_1}$
Total	0	0	$F(T_2) - F(T_1) \times (1+r)^{T_2-T_1} + FV(D)$

(a) *If the bi-annual current interest rate is 2.5 %, and the bi-annual dividend rate of the index is 1.5 %, is there an arbitrage opportunity available? If there is, calculate the profits available on the strategy.*
(b) *Is there an arbitrage opportunity if the interest rate that can be earned on the proceeds of a short sale is only 2 % bi-annually?*

(a) The bi-annual net cost of carry is $1+r-d = 1+0.025-0.015 = 1.01 = 1\,\%$. The detailed cash flow can be seen in Table 2.6.

Thus, the arbitrage profit is 12.

(b) Now consider a lower bi-annual interest rate of 2 %. From Table 2.7 which displays the detailed cash flow, we could see the arbitrage opportunity has gone.

Exercise 2.10 (Hedging Strategy). *A portfolio manager holds a portfolio that mimics the S&P 500 index. The S&P 500 index started at the beginning of this year at* 800 *and is currently at* 923.33. *The December S&P 500 futures price is currently* 933.33 *USD. The manager's fund was valued at ten million USD at the beginning of this year. Since the fund has already generated a handsome return last year, the manager wishes to lock in its current value. That is, the manager is willing to give up potential increases in order to ensure that the value of the fund does not decrease. How can you lock in the value of the fund implied by the December futures contract? Show that the hedge does work by considering the value of your net hedged position when the S&P 500 index finishes the year at* 833.33 *and* 1,000 *USD.*

First note that at the December futures price of 933.33 USD, the return on the index, since the beginning of the year, is $933.33/800 - 1 = 16.7\,\%$. If the manager is able to lock in this return on the fund, the value of the fund will be $1.1667 \cdot 10 = 11.67$ million USD. Since the notional amount underlying the S&P 500 futures contract is $500 \cdot 933.33 = 466,665$ USD, the manager can lock in the 16.67 % return by selling $11{,}666{,}625/466{,}665 = 25$ contracts.

Suppose the value of the S&P 500 index is 833.33 at the end of December. The value of the fund will be $833.33/800 \cdot 10 = 10.42$ million USD. The gain on the futures position will be $-25 \cdot 500(833.33 - 933.33) = 1.25$ million USD. Hence,

Table 2.6 Cash flow table for this strategy

Action	CF_0	CF_T
Buy futures contract	0	S_T−3,220
Sell stock	3,200	$-S_T - 0.015 \cdot 3{,}200$
Lend proceeds of sale	−3,200	3,200 ·1.025
	0	12

Table 2.7 Cash flow table with a lower interest rate

Action	CF_0	CF_T
Buy futures contract	0	S_T− 3,220
Sell stock	3,200	$S_T - 0.015 \cdot 3{,}200$
Lend proceeds of sale	−3,200	$3,200 \cdot 1.02$
	0	−4

the total value of the hedged position is $10.42 + 1.25 = 11.67$ million USD, locking in the 16.67 % return for the year.

Now suppose that the value of the S&P 500 index is 1,000 at the end of December. The value of the fund will be $1{,}000/800 \cdot 10 = 12.5$ million USD. The gain on the futures position will be $-25 \cdot 500(1{,}000 - 933.33) = -0.83$ million USD. Hence, the total value of the hedged position is $12.5 - 0.83 = 11.67$ million USD, again locking in the 16.67 % return for the year.

Exercise 2.11 (Forward Exchange Rate). *The present exchange rate between the USD and the EUR is* 1.22 *USD/EUR. The price of a domestic* 180*-day Treasury bill is* 99.48 *USD per* 100 *USD face value. The price of the analogous EUR instrument is* 99.46 *EUR per* 100 *EUR face value.*

(a) What is the theoretical 180*-day forward exchange rate?*
(b) Suppose the 180*-day forward exchange rate available in the marketplace is* 1.21 *USD/EUR. This is less than the theoretical forward exchange rate, so an arbitrage is possible. Describe a risk-free strategy for making money in this market. How much does it gain, for a contract size of* 100 *EUR?*

(a) The theoretical forward exchange rate is

$$1.22 \cdot 0.9946/0.9948 = 1.2198 \text{ USD/EUR}.$$

(b) The price of the forward is too low, so the arbitrage involves buying forwards. Firstly, go long on a forward contract for 100 EUR with delivery price 1.21 USD/EUR. Secondly, borrow $\exp(-qT)$ EUR now, convert to dollars at 1.22 USD/EUR and invest at the dollar rate.

At maturity, fulfill the contract, pay $1.21 \cdot 100$ USD for 100 EUR, and clear your cash positions. You have $(1.2198 - 1.21) \cdot 100 = 0.0098 \cdot 100$ USD. That is, you make 0.98 USD at maturity risk-free.

Table 2.8 Cash flow table for zero-net-investment arbitrage portfolio

Action	CF_0	CF_T
Short shares	2,500	$-(S_T + 40)$
Long futures	0	S_T−2,530
Long zero-bonds	−2,500	2,576.14
Total	0	6.14

Exercise 2.12 (Valuation of a Forward Contract). *What is the value of a forward contract with $K = 100$, $S_t = 95$, $r = 10\,\%$, $d = 5\,\%$ and $\tau = 0.5$?*

The payoff of the forward contract can be duplicated with buying $\exp(-d\tau)$ stocks and short selling zero bonds with nominal value $K\exp(-r\tau)$. So

$$\begin{aligned} V_{K,T}(S_t,\tau) &= \exp(-0.05 \cdot 0.5) \cdot 95 - 100 \cdot \exp(-0.10 \cdot 0.5) \\ &= -2.4685 \end{aligned}$$

Thus, the buyer of the forward contract should be paid 2.4685 for this deal.

Exercise 2.13 (Put-Call Parity). *Suppose there is a 1-year future on a stock-index portfolio with the future price 2,530 USD. The current stock index is 2,500, and a 2,500 USD investment in the index portfolio will pay a year-end dividend of* 40 *USD. Assume that the 1-year risk-free interest rate is 3 %.*

(a) Is this future contract mispriced?
(b) If there is an arbitrage opportunity, how can an investor exploit it using a zero-net investment arbitrage portfolio?
(c) If the proceeds from the short sale of the shares are kept by the broker (you do not receive interest income from the fund), does this arbitrage opportunity still exist?
(d) Given the short sale rules, how high and how low can the futures price be without arbitrage opportunities?

(a) The price of a future can be found as follows:

$$\begin{aligned} F_0 &= S_0 \exp(r\tau) - D \\ &= 2500 \cdot \exp(0.03) - 40 \\ &= 2576.14 - 40 \\ &= 2536.14 > 2530 \end{aligned}$$

This shows that the future is priced 6.14 EUR lower.

(b) Zero-net-investment arbitrage portfolio Cash flow for this portfolio is described in Table 2.8.

Table 2.9 Cash flow table for the no interest income case

Action	CF_0	CF_T
Short shares	2,500	$-(S_T + 40)$
Long futures	0	$S_T - 2,530$
Long zero-bonds	−2,500	2,500
Total	0	−70

Table 2.10 Cash flow table for this strategy

Action	CF_0	CF_T
Short shares	2,500	$-(S_T + 40)$
Long futures	0	$S_T - F_0$
Long zero-bonds	−2,500	2,500
Total	0	2,460$-F_0$

(c) No interest income case

According to Table 2.9, the arbitrage opportunity does not exist.

(d) To avoid arbitrage, 2,460$-F_0$ must be non-positive, so $F_0 \geq$ 2,460. On the other hand, if F_0 is higher than 2,536.14, an opposite arbitrage opportunity (buy stocks, sell futures) opens up. Finally we get the no-arbitrage band 2,460$\leq F_0 \leq$ 2,536.14 (Table 2.10)

Exercise 2.14 (Hedging Strategy). *The price of a stock is* 50 *USD at time* $t = 0$. *It is estimated that the price will be either* 25 *or* 100 *USD at* $t = 1$ *with no dividends paid. A European call with an exercise price of* 50 *USD is worth* C *at time* $t = 0$. *This call will expire at time* $t = 1$. *The market interest rate is 25 %.*

(a) What return can the owner of the following hedge portfolio expect at $t = 1$ *for the following actions: sell 3 calls for* C *each, buy* 2 *stocks for* 50 *USD each and borrow* 40 *USD at the market interest rate*
(b) Calculate the price C *of a call.*

(a) By setting up a portfolio where 3 calls are sold $3C$, 2 stocks are bought $-2 \cdot 50$ and 40 USD are borrowed at the market interest rate at the current time t, the realised immediate profit is $3C - 60$. The price of the call option can be interpreted as the premium to insure the stocks against falling below 50 USD. At time $t = 1$, if the price of the stock is less than the exercise price ($S_t < K$) the holder does not exercise the call options, otherwise he does. When the price of the stock at time $t = 1$ is equal to 25 USD, the holder does not exercise the call option, but he does when the price of the stock at time $t = 1$ is 100 USD. Also at time $t = 1$, the holder gets the value $2S_1$ by purchasing two stocks at $t = 0$ and pays back the borrowed money at the interest rate of 25 %. The difference of the value of the portfolio with the corresponding stock price 25 USD or 100 USD at $t = 1$ is shown in Table 2.11. At time $t = 1$, the cash flow is independent of

Table 2.11 Portfolio value at time $t = 1$ of Exercise 2.14

Action	CF_0	$CF_1(S_1 = 25)$	$CF_1(S_1 = 100)$
Sell 3 calls	$3C$	0	$-3(100 - 50) = -150$
Buy 2 stocks	$-2 \cdot 50 = -100$	$2 \cdot 25 = 50$	$2 \cdot 100 = 200$
Borrow	$40n$	$-40(1 + 0.25) = -50$	-50
Total	$3C - 60$	0	0

the stock, which denotes this strategy as risk-free. That is, the owner does not expect any return from the described hedge portfolio.

(b) The price of the call of this hedge portfolio is equal to the present value of the cash flows at $t = 1$ minus the cash flow at $t = 0$. In this case we have that the present value of cash flows at $t = 1$ is equal to zero and the cash flow at time $t = 0$ is $3C - 60$. Therefore, the value of the call option is equal to $C = 20$. Here the martingale property is verified, since the conditional expected value of the stock price at time $t = 1$, given the stock prices up to time $t = 0$, is equal to the value at the earlier time $t = 0$.

Chapter 3
Basic Concepts of Probability Theory

概率论基本概念

不入虎穴，焉得虎子. 后汉书·班超传

If you don't enter the tiger's den, how can you catch the tiger's cub?

The Book of the Later Han: the Biography of Ban Chao

This part is an introduction to standard concepts of probability theory. We discuss a variety of exercises on moment and dependence calculations with a real marketing example. We also study the characteristics of transformed random vectors, e.g. distributions and various statistical measures. Another feature that needs to be considered is various conditional statistical measures and their relations with corresponding marginal and joint distributions. Two more exercises are given in order to distinguish the differences between numerical statistic measures and statistical properties.

Exercise 3.1 (χ^2 distribution). *If $X \sim N(0,1)$ then X^2 has χ^2_1 distribution with a pdf shown in Fig. 3.1. Calculate the distribution function and the density of the χ^2_1 distribution.*

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_3,

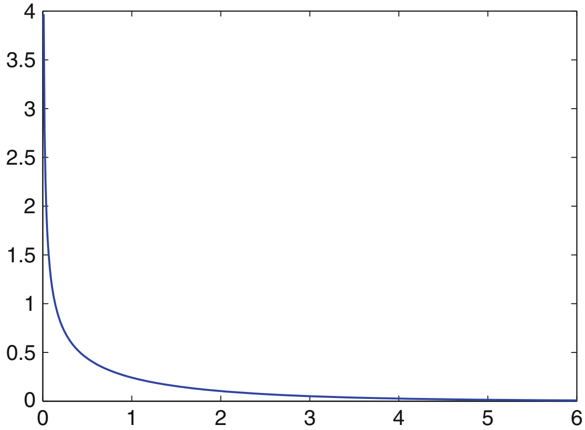


Fig. 3.1 Pdf of a χ^2_1 distribution `SFSchisq`

For $t > 0$

$$
\begin{aligned}
\mathrm{P}(X^2 \le t) &= \mathrm{P}(-\sqrt{t} \le X \le \sqrt{t}) = \Phi(\sqrt{t}) - \Phi(-\sqrt{t}) \\
&= 2\Phi(\sqrt{t}) - 1 = 2\left(\int_{-\infty}^{0} \varphi(x)dx + \int_{0}^{\sqrt{t}} \varphi(x)dx\right) - 1 \\
&= 2\left(\frac{1}{2} + \frac{1}{\sqrt{2\pi}}\int_0^{t/2} (2z)^{-1/2} e^{-z} dz\right) - 1 = \frac{1}{\sqrt{\pi}}\int_0^{t/2} z^{1/2-1} e^{-z} dz
\end{aligned}
$$

The function $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx$ is called *gamma function* and has the following properties: $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(t+1) = t\Gamma(t)$. The lower incomplete gamma function is defined by $\gamma(a,t) = \int_0^t x^{a-1}e^{-x}dx$. Therefore the cdf of χ^2_1 can be expressed as $\gamma(1/2, t/2)/\Gamma(1/2)$.

To calculate the density one takes the derivative with respect to the upper limit of the integral, which yields $f(t) = \{\Gamma(1/2)2te^t\}^{-1/2}$

Exercise 3.2 (χ^2 distribution). *If $X_1, \dots, X_n$ are i.i.d. $\sim N(0,1)$ then $\sum_{i=1}^n X_i^2$ has χ^2_n distribution with pdf as in Fig. 3.2. Calculate mean and variance of the χ^2_n distribution.*

As the second and fourth moments of the standard normal distribution are 1 and 3 correspondingly, we have:

$$
\mathsf{E}(\sum_{i=1}^n X_i^2) = n\mathsf{E}X_1^2 = n
$$

$$
\mathsf{Var}(\sum_{i=1}^n X_i^2) = n\mathsf{Var}X_1^2 = n(\mathsf{E}X_1^4 - \mathsf{E}X_1^2) = n(3-1) = 2n
$$

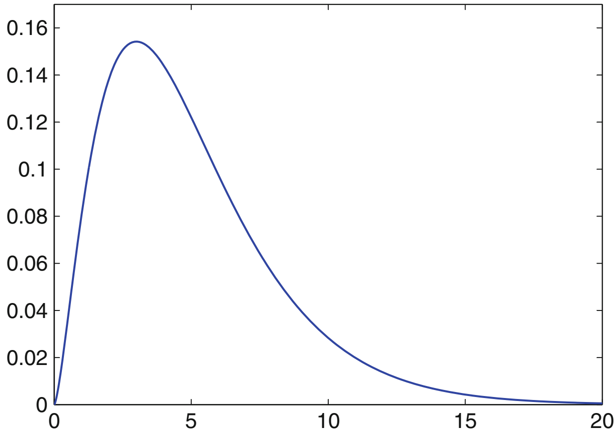


Fig. 3.2 Pdf of a χ^2_5 distribution SFSchisq

Exercise 3.3 (Random Walk). *Check that the random variable* X *with* $\mathrm{P}(X=1) = 1/2$, $\mathrm{P}(X=-4) = 1/3$, $\mathrm{P}(X=5) = 1/6$ *has skewness* 0 *but is not distributed symmetrically.*

$$\mu = \mathsf{E}(X) = 1 \cdot 1/2 + (-4) \cdot 1/3 + 5 \cdot 1/6 = 0$$

$\mathsf{E}(X-\mu)^3 = 1 \cdot 1/2 + (-4)^3 \cdot 1/3 + 5^3 \cdot 1/6 = 0$, which implies that its skewness $\mathsf{E}(X-\mu)^3/\sigma^3$ is 0. It is easy to see that the random variable is not distributed symmetrically.

Exercise 3.4 (Independence). *Show that if* $\mathsf{Cov}(X,Y) = 0$ *it does not imply that* X *and* Y *are independent.*

Consider a standard normal random variable X and a random variable $Y = X^2$, which is not independent of X. Here we have

$$\mathsf{Cov}(X,Y) = \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y) = \mathsf{E}(X^3) = 0.$$

Exercise 3.5 (Correlation). *Show that the correlation is invariant w.r.t. linear transformations.*

Since $\mathsf{Corr}(X,Y) = \mathsf{Corr}(Y,X)$, it suffices to show $\mathsf{Corr}(aX+b,Y) = \mathsf{Corr}(X,Y)$ since then $\mathsf{Corr}(aX+b,cY+d) = \mathsf{Corr}(X,cY+d) = \mathsf{Corr}(X,Y)$ for $a > 0$. From the definition

$$\begin{aligned}\mathsf{Corr}(aX+b,Y) &= [\mathsf{E}\{(aX+b)Y\} - \mathsf{E}(aX+b)\mathsf{E}(Y)]/\sigma(aX+b)\sigma(Y)\\ &= a\mathsf{Cov}(X,Y)/a\sigma(X)\sigma(Y) = \mathsf{Corr}(X,Y).\end{aligned}$$

Note that the correlation does not need to be invariant to nonlinear transformations, see Exercise 3.4. More generally, if $Y = X^n$, we have

$$\operatorname{Corr}(X, Y) = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}$$

$$= \frac{\mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y)}{\sqrt{\operatorname{Var}(X)\{\mathsf{E}(Y^2) - \mathsf{E}^2(Y)\}}} = \frac{\mathsf{E}(X^{n+1}) - \mathsf{E}(X)\mathsf{E}(X^n)}{\sqrt{\{\mathsf{E}(X^2) - \mathsf{E}^2(X)\}\{\mathsf{E}(X^{2n}) - \mathsf{E}^2(X^n)\}}}$$

So we could see that $\operatorname{Corr}(X, X)$ is not always equal to $\operatorname{Corr}(X, Y)$ in general. Thus correlation is not always invariant under nonlinear transformations.

Exercise 3.6 (Independence). *Let* $\{X_i\}_{i=1}^n \overset{i.i.d.}{\sim} N(\mu, \sigma)$. *Show that the random variable* $\overline{X}$ *and* $X_i - \overline{X}$ *are independent for all i.*

Since both variables are normal, it is enough for independence to show that they are uncorrelated.

$$\operatorname{Cov}(\overline{X}, X_i - \overline{X}) = \mathsf{E}[\overline{X}(X_i - \overline{X})]$$

Since

$$\mathsf{E}[X_i - \overline{X}] = \mathsf{E}[-\frac{1}{n}X_1 - \ldots + (1 - \frac{1}{n})X_i - \ldots - \frac{1}{n}X_n]$$
$$= (-\frac{n-1}{n})\mu + (1 - \frac{1}{n})\mu = 0$$

But

$$\mathsf{E}[(\frac{1}{n}X_1 + \ldots + \frac{1}{n}X_i + \ldots + \frac{1}{n}X_n)X_i] = \frac{n-1}{n}\mathsf{E}[X_1X_i] + \frac{1}{n}\mathsf{E}[X_i^2]$$
$$= \frac{n-1}{n}\mu^2 + \frac{1}{n}(\sigma^2 + \mu^2)$$
$$= \mu^2 + \frac{\sigma^2}{n}$$

and from $\operatorname{Var}[\overline{X}] = \mathsf{E}[\overline{X}^2] - (\mathsf{E}[\overline{X}])^2$ we get $\mathsf{E}[\overline{X}^2] = \frac{\sigma^2}{n} + \mu^2$.

Then $\mathsf{E}[\overline{X}X_i - \overline{X}^2] = \mu^2 + \frac{\sigma^2}{n} - (\frac{\sigma^2}{n} + \mu^2) = 0$ and therefore $\operatorname{Cov}(\overline{X}, X_i - \overline{X}) = 0$.

Exercise 3.7 (Correlation). *We consider a bivariate exchange rates example, two European currencies, EUR and GBP, with respect to the USD. The sample period is 01/01/2002–01/01/2009 with altogether n = 1,828 observations. Figure 3.3 shows the time series of returns on both exchange rates.*

Compute the correlation of the two exchange rate time series and comment on the sign of the correlation.

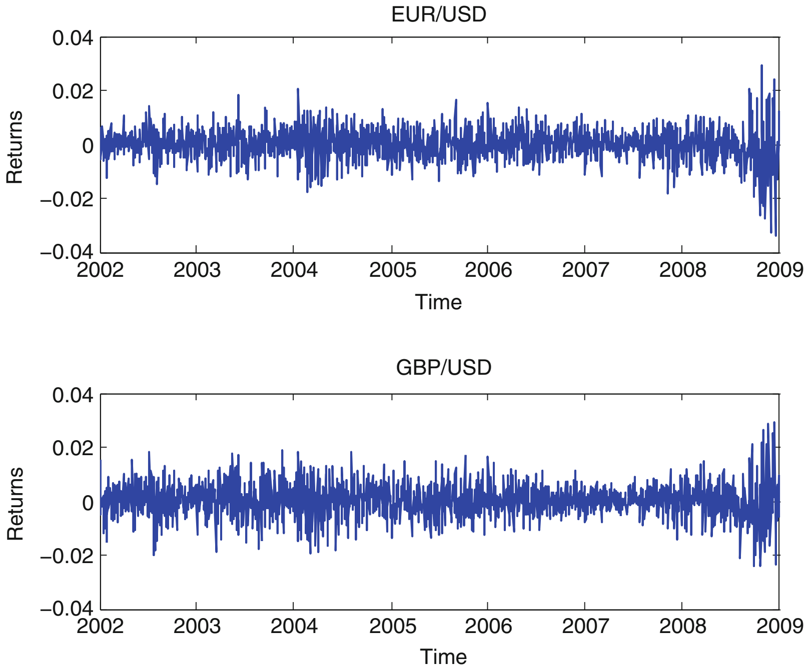


Fig. 3.3 Exchange rate returns. `SFSmvol01`

The correlation $r = 0.7224$ `SFSmvol01` says that the relationship between EUR/USD and GBP/USD exchange rates is positive as predicted by the economic theory. This also confirms our intuition of mutual dependence in exchange markets.

Exercise 3.8 (Conditional Moments). *Compute the conditional moments* $\mathsf{E}(X_2 \mid x_1)$ *and* $\mathsf{E}(X_1 \mid x_2)$ *for the pdf of*

$$f(x_1, x_2) = \begin{cases} \frac{1}{2}x_1 + \frac{3}{2}x_2 & 0 \le x_1, x_2 \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

The marginal densities of X_1 and X_2, for $0 \le x_1, x_2 \le 1$, are

$$f_{X_1}(x_1) = \int_0^1 f(x_1, x_2)dx_2 = \left[\frac{1}{2}x_1x_2 + \frac{3}{4}x_2^2\right]_0^1 = \frac{1}{2}x_1 + \frac{3}{4}$$

and

$$f_{X_2}(x_2) = \int_0^1 f(x_1, x_2)dx_1 = \left[\frac{1}{4}x_1^2 + \frac{3}{2}x_1x_2\right]_0^1 = \frac{1}{4} + \frac{3}{2}x_2.$$

Now, the conditional expectations, for $0 \le x_1, x_2 \le 1$, can be calculated as follows:

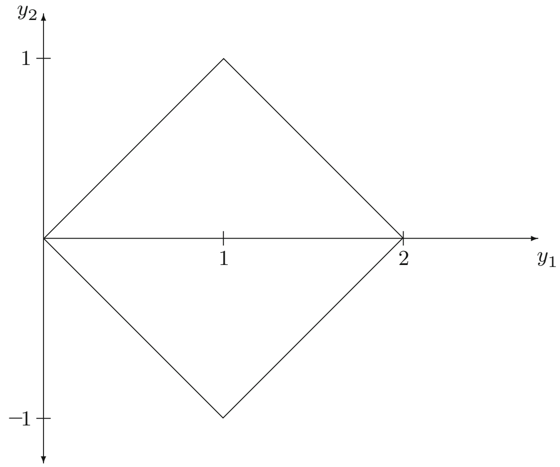


Fig. 3.4 The support of the pdf $f_Y(y_1, y_2)$ given in Exercise 3.9

$$\mathsf{E}(X_2|x_1) = \int_0^1 x_2 f(x_2|x_1) dx_2 = \int_0^1 x_2 \left(\frac{\frac{1}{2}x_1 + \frac{3}{2}x_2}{\frac{1}{2}x_1 + \frac{3}{4}} \right) dx_2$$

$$= \left[\frac{\frac{x_1 x_2^2}{4} + \frac{x_2^3}{2}}{\frac{3}{4} + \frac{x_1}{2}} \right]_0^1 = \frac{x_1 + 2}{3 + 2x_1}$$

$$\mathsf{E}(X_1|x_2) = \int_0^1 x_1 f(x_1|x_2) dx_1 \int_0^1 x_1 \left(\frac{\frac{1}{2}x_1 + \frac{3}{2}x_2}{\frac{3}{2}x_2 + \frac{1}{4}} \right) dx_1$$

$$= \left[\frac{\frac{x_1^3}{6} + \frac{3x_1^2 x_2}{4}}{\frac{1}{4} + \frac{3x_2}{2}} \right]_0^1 = \frac{2 + 9x_2}{3 + 18x_2}$$

Exercise 3.9 (Probability Density Function). *Show that the function*

$$f_Y(y_1, y_2) = \begin{cases} \frac{1}{2}y_1 - \frac{1}{4}y_2 & 0 \le y_1 \le 2, \ |y_2| \le 1 - |1 - y_1| \\ 0 & \textit{otherwise} \end{cases}$$

is a probability density function.

The area for which the above function is non-zero is plotted in Fig. 3.4.

In order to verify that $f_Y(y_1, y_2)$ is a two-dimensional pdf, we have to check that it is nonnegative and that it integrates to 1.

It is easy to see that the function $f_Y(y)$ is nonnegative inside the square plotted in Fig. 3.4 since $y_1 \ge 0$ and $y_1 \ge y_2$ implies that $y_1/2 - y_2/4 > 0$.

It remains to verify that the function $f_Y(y)$ integrates to one by calculating the integral

$$\int f_Y(y)dy$$

for which we easily obtain the following:

$$\begin{aligned}
\int\int f_Y(y_1,y_2)dy_2,y_1 &= \int_0^1\int_{-y_1}^{y_1} f_Y(y)dy_2dy_1 + \int_1^2\int_{y_1-2}^{2-y_1} f_Y(y)dy_2dy_1\\
&= \int_0^1\int_{-y_1}^{y_1} \frac{1}{2}y_1 - \frac{1}{4}y_2dy_2dy_1 + \int_1^2\int_{y_1-2}^{2-y_1} \frac{1}{2}y_1 - \frac{1}{4}y_2dy_2dy_1
\end{aligned}$$

$$\begin{aligned}
&= \int_0^1 \left[\frac{1}{2}y_1y_2 - \frac{1}{8}y_2^2\right]_{-y_1}^{y_1} dy_1 + \int_1^2 \left[\frac{1}{2}y_1y_2 - \frac{1}{8}y_2^2\right]_{y_1-2}^{2-y_1} dy_1\\
&= \int_0^1 y_1^2dy_1 + \int_1^2 -y_1^2 + 2y_1dy_1\\
&= \left[\frac{1}{3}y_1^3\right]_0^1 + \left[-\frac{1}{3}y_1^3 + y_1^2\right]_1^2 = \frac{1}{3} + \frac{2}{3} = 1.
\end{aligned}$$

Exercise 3.10 (Conditional Expectation). *Prove that* $\mathsf{E}X_2 = \mathsf{E}\{\mathsf{E}(X_2|X_1)\}$, *where* $\mathsf{E}(X_2|X_1)$ *is the conditional expectation of* X_2 *given* X_1.

Since $\mathsf{E}(X_2|X_1 = x_1)$ is a function of x_1, it is clear that $\mathsf{E}(X_2|X_1)$ is a random vector (function of random vector X_1).

Assume that the random vector $X = (X_1, X_2)^\top$ has the density $f(x_1, x_2)$. Then

$$\begin{aligned}
\mathsf{E}\{\mathsf{E}(X_2|X_1)\} &= \int\left\{\int x_2 f(x_2|x_1)dx_2\right\} f(x_1)dx_1\\
&= \int\left\{\int x_2 \frac{f(x_2,x_1)}{f(x_1)}dx_2\right\} f(x_1)dx_1 = \int\int x_2 f(x_2,x_1)dx_2dx_1\\
&= \mathsf{E}X_2.
\end{aligned}$$

Exercise 3.11 (Conditional Variance). *The conditional variance is defined as* $\mathsf{Var}(Y|X) = \mathsf{E}[\{Y - \mathsf{E}(Y|X)\}^2|X]$. *Show that* $\mathsf{Var}(Y) = \mathsf{E}\{\mathsf{Var}(Y|X)\} + \mathsf{Var}\{\mathsf{E}(Y|X)\}$.

$$
\begin{aligned}
\mathsf{E}\{\mathsf{Var}(Y|X)\} &= \mathsf{E}(\mathsf{E}[\{Y-\mathsf{E}(Y|X)\}^2|X]) \\
&= \mathsf{E}[\mathsf{E}(Y^2|X) - 2\mathsf{E}\{Y\mathsf{E}(Y|X)|X\} + \mathsf{E}\{\mathsf{E}(Y|X)^2|X\}] \\
&= \mathsf{E}\{\mathsf{E}(Y^2|X)\} - 2\mathsf{E}\{\mathsf{E}(Y|X)\mathsf{E}(Y|X)|X\} + \mathsf{E}[\mathsf{E}\{\mathsf{E}(Y|X)^2|X\}] \\
&= \mathsf{E}\{Y^2 - \mathsf{E}(Y|X)\mathsf{E}(Y|X)\} \qquad (3.1) \\
\mathsf{Var}\{\mathsf{E}(Y|X)\} &= \mathsf{E}([\mathsf{E}(Y|X) - \mathsf{E}\{\mathsf{E}(Y|X)\}]^2) \\
&= \mathsf{E}\{\mathsf{E}(Y|X)\mathsf{E}(Y|X)\} - 2\mathsf{E}\{\mathsf{E}(Y|X)E(Y)\} + E(Y)^2 \\
&= \mathsf{E}\{\mathsf{E}(Y|X)\mathsf{E}(Y|X)\} - \mathsf{E}^2(Y). \qquad (3.2)
\end{aligned}
$$

Summing up (3.1) and (3.2) yields $\mathsf{E}\{\mathsf{Var}(Y|X)\} + \mathsf{Var}\{\mathsf{E}(Y|X)\} = \mathsf{E}(Y^2) - \mathsf{E}^2(Y) = \mathsf{Var}(Y)$.

Exercise 3.12 (Marginal Distribution). *Consider the pdf*

$$f(x_1, x_2) = \frac{1}{8x_2}\exp\{-(x_1/2x_2 + x_2/4)\} \qquad x_1, x_2 > 0.$$

Compute $f(x_2)$ and $f(x_1|x_2)$.

The marginal distribution of x_2 can be calculated by integrating out x_1 from the joint pdf $f(x_1, x_2)$:

$$
\begin{aligned}
f_{X_2}(x_2) &= \int_0^{+\infty} f(x_1, x_2)dx_1 \\
&= -\frac{1}{4}\exp(-x_2/4)\int_0^{+\infty} -1/(2x_2)\exp(-x_1/2x_2)\,dx_1 \\
&= \frac{1}{4}\exp(-x_2/4)\,[\exp(-x_1)]_0^{+\infty} \\
&= \frac{1}{4}\exp(-x_2/4)
\end{aligned}
$$

for $x_2 > 0$, in other words, the distribution of X_2 is exponential with expected value $\mathsf{E}(X_2) = 4$.

The conditional distribution $f(x_1|x_2)$ is calculated as a ratio of the joint pdf $f(x_1, x_2)$ and the marginal pdf $f_{X_2}(x_2)$:

$$
\begin{aligned}
f_{X_1|X_2=x_2}(x_1) &= f(x_1, x_2)/f_{X_2}(x_2) \\
&= \exp(-x_1/2x_2)/(2x_2),
\end{aligned}
$$

for $x_1, x_2 > 0$. Note that this is just the exponential distribution with expected value $2x_2$.

Exercise 3.13 (Asymptotic Distribution). *A European car manufacturer has tested a new model and reports on the consumption of gasoline (X_1) and oil (X_2). The expected consumption of gasoline is 8 L per 100 km (μ_1) and the expected consumption of oil is 1 L per 10.000 km (μ_2). The measured consumption of gasoline is 8.1 L per 100 km ($\overline{x}_1$) and the measured consumption of oil is 1.1 L per 10,000 km ($\overline{x}_2$). The asymptotic distribution of*

$$\sqrt{n}\left\{\begin{pmatrix}\overline{x}_1\\ \overline{x}_2\end{pmatrix} - \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\right\} \text{ is } N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0.1 & 0.05\\ 0.05 & 0.1\end{pmatrix}\right).$$

For the American market the basic measuring units are miles (1 mile ≈ 1.6 km) and gallons (1 gallon ≈ 3.8 L). The consumptions of gasoline (Y_1) and oil (Y_2) are usually reported in miles per gallon. Can you express $\overline{y}_1$ and $\overline{y}_2$ in terms of $\overline{x}_1$ and $\overline{x}_2$? Recompute the asymptotic distribution for the American market.

The transformation of "liters per 100 km" to "miles per gallon" is given by the function

$$\begin{aligned} x \text{ liters per 100 km } &= 1.6x/380 \text{ gallons per mile}\\ &= 380/(1.6x) \text{ miles per gallon.}\end{aligned}$$

Similarly, we transform the oil consumption

$$x \text{ liters per 10000 km } = 38000/(1.6x) \text{ miles per gallon.}$$

Thus, the transformation is given by the functions

$$\begin{aligned} f_1(x) &= 380/(1.6x)\\ f_2(x) &= 38000/(1.6x).\end{aligned}$$

According to Härdle and Simar (2012, Theorem 4.11), the asymptotic distribution is

$$\sqrt{n}\left\{\begin{pmatrix}f_1(\overline{x}_1)\\ f_2(\overline{x}_2)\end{pmatrix} - \begin{pmatrix}f_1(\mu_1)\\ f_2(\mu_2)\end{pmatrix}\right\} \sim \mathrm{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \mathcal{D}^\top \begin{pmatrix}0.1 & 0.05\\ 0.05 & 0.1\end{pmatrix}\mathcal{D}\right),$$

where

$$D = \left(\frac{\partial f_j}{\partial x_i}\right)(x)\bigg|_{x=\mu}$$

is the matrix of all partial derivatives. In our example,

$$\begin{aligned}
\mathcal{D} &= \begin{pmatrix} -\frac{380}{1.6x_1^2} & 0 \\ 0 & -\frac{38000}{1.6x_2^2} \end{pmatrix}\Bigg|_{x=\mu} \\
&= \begin{pmatrix} -\frac{380}{1.6\overline{x}_1^2} & 0 \\ 0 & -\frac{38000}{1.6x_2^2} \end{pmatrix} \\
&= \begin{pmatrix} -3.62 & 0 \\ 0 & -19628.10 \end{pmatrix}.
\end{aligned}$$

Hence, the variance of the transformed random variable Y is given by

$$\begin{aligned}
\Sigma_Y &= \mathcal{D}^\top \begin{pmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{pmatrix} \mathcal{D} \\
&= \begin{pmatrix} -3.62 & 0 \\ 0 & -19628.10 \end{pmatrix} \begin{pmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{pmatrix} \begin{pmatrix} -3.62 & 0 \\ 0 & -19628.10 \end{pmatrix} \\
&= \begin{pmatrix} 1.31 & 3552.69 \\ 3552.69 & 38526230.96 \end{pmatrix}.
\end{aligned}$$

The average fuel consumption, transformed to American units of measurements is $\overline{y}_1 = 29.32$ miles per gallon and the transformed oil consumption is $\overline{y}_2 = 19{,}628.10$. The asymptotic distribution is

$$\sqrt{n}\left\{\begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} - \begin{pmatrix} f_1(\mu_1) \\ f_2(\mu_2) \end{pmatrix}\right\} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.31 & 3552.69 \\ 3552.69 & 38526230.96 \end{pmatrix}\right).$$

Chapter 4
Stochastic Processes in Discrete Time

Processus artis coniectandi, qui spatio temporis discreto fiunt
Vitam regit fortuna, non sapientia.
Fortune, not wisdom, rules lives.

Marcus Tullius Cicero, Tusculanarum Disputationum LIX

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 $X_0 = 0$. This means that even if the starting point is known, there are many possible routes the process might take, some of them with a higher probability. In this section, we exclusively consider processes in *discrete time*, i.e. processes which are observed at equally spaced points of time $t = 0, 1, 2, \ldots$. In other words, a discrete process is considered to be an approximation of the continuous counterpart. Hence, it is important to start with discrete processes in order to understand sophisticated continuous processes. In particular, a Brownian motion is a limit of random walks and a stochastic differential equation is a limit of stochastic difference equations. A random walk is a stochastic process with independent, identically distributed binomial random variables which can serve as the basis for many stochastic processes.

Typical examples are daily, monthly or yearly observed economic data as stock prices, rates of unemployment or sales figures.

In order to get an impression of stochastic processes in discrete time, we plot the time series for the Coca-Cola stock price. The results are displayed in Fig. 4.1. If prices do not vary continuously, at least they vary frequently, and the stochastic process has thus proved its usefulness as an approximation of reality.

Exercise 4.1 (Geometric Brownian Motion). *Construct a simulation for a random stock price movement in discrete time with the characteristics given in Table 4.1 from a geometric Brownian motion.*

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_4, © Springer-Verlag Berlin Heidelberg 2013

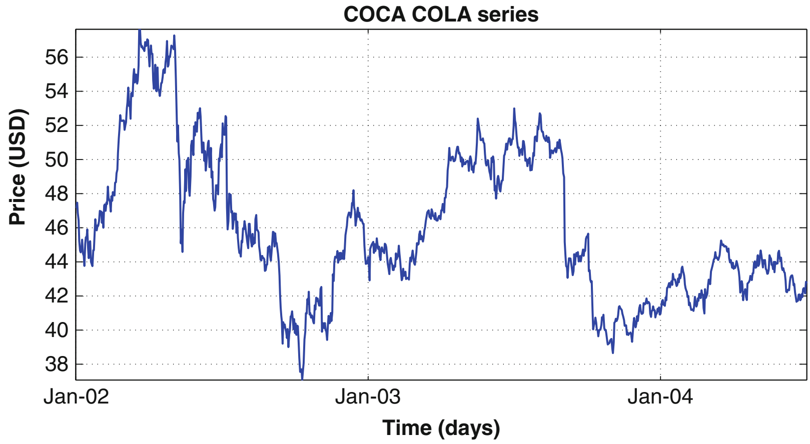


Fig. 4.1 Stock price of Coca-Cola

Table 4.1 Characteristics to simulate a random stock price movement in discrete time

Default values	
Initial stock price S_0	49
Initial time	0
Time to maturity T	20 weeks
Time interval Δt	1 week
Volatility σ p.a.	0.20
Expected return μ p.a.	0.13

The numerical procedure to simulate the stock price movement in discrete time with characteristics described in Table 4.1 is given by defining the process $S_i = S_{i-1}\exp\{X_i\sigma\sqrt{T/n} + (\mu - \sigma^2/2)T/n\}$, with $i = 0, \ldots, n$ where n denotes the number of time intervals, $\Delta t = T/n$ and $X \sim \mathrm{N}(0,1)$ denotes a standard normal r.v. Fig. 4.2 displays the simulation of a random stock price movement in discrete time with $\Delta t = 1$ week and 1 day respectively.

Exercise 4.2 (Random Walk). *Consider an ordinary random walk* $X_t = \sum_{k=1}^{t} Z_k$ *for* $t = 1, 2, \ldots,$ $X_0 = 0$, *where* $Z_1, Z_2, \ldots$ *are i.i.d. with* $\mathrm{P}(Z_k = 1) = p$ *and* $\mathrm{P}(Z_k = -1) = 1 - p$, $p \in (0,1)$. *Calculate*

(a) $\mathrm{P}(X_t > 0)$
(b) $\mathrm{P}(X_t = 1)$
(c) $\mathrm{P}(Z_2 = 1|X_3 = 1)$

(a) Let $Y_k = (Z_k + 1)/2$ then $B_t = \sum_{k=1}^{t} Y_k$ has binomial distribution $B(t,p)$. It is easy to see that $X_t = 2B_t - t$.

$$\mathrm{P}(X_t > 0) = \mathrm{P}(2B_t - t > 0) = \mathrm{P}(B_t > t/2) = \sum_{k=\lceil t/2+1\rceil}^{t} \binom{t}{k} p^k (1-p)^{t-k},$$

where $\lceil x \rceil$ means the largest integer not larger than x.

(b) $\mathrm{P}(X_t = 1) = \mathrm{P}(2B_t - t = 1) = \mathrm{P}\{B_t = (1+t)/2\}$

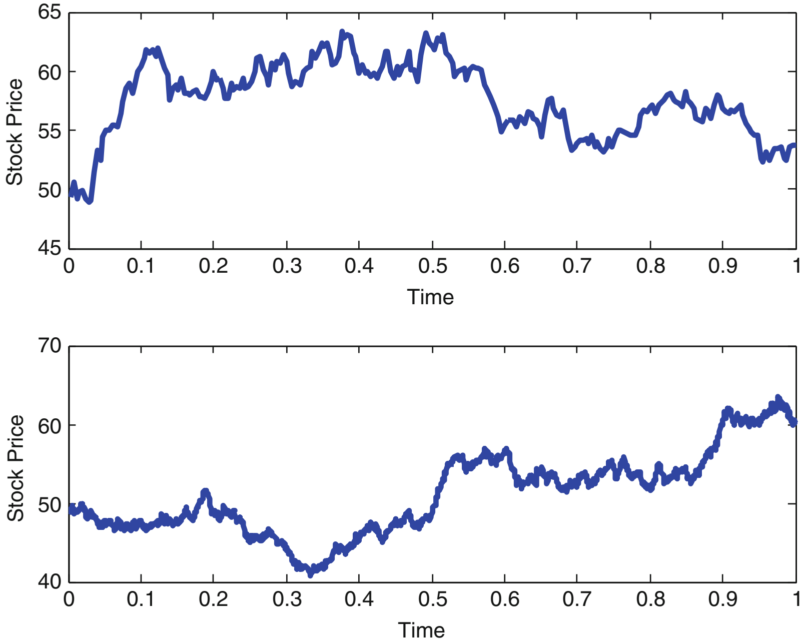


Fig. 4.2 Simulation of a random stock price movement in discrete time with $\Delta t = 1$ day (*up*) and 1 (*down*) week respectively. `SFSrwdiscretetime`

If t is even $\mathrm{P}(X_t = 1) = 0$ and in case t is odd

$$\mathrm{P}(X_t = 1) = \binom{t}{\frac{1+t}{2}} p^{\frac{1+t}{2}}(1-p)^{\frac{t-1}{2}}$$

In particular $\mathrm{P}(X_3 = 1) = 3p^2(1-p)$.

(c)

$$\begin{aligned}\mathrm{P}(Z_2 = 1|X_3 = 1) &= \mathrm{P}(Z_2 = 1 \wedge X_3 = 1)/\mathrm{P}(X_3 = 1)\\ &= \mathrm{P}(Z_2 = 1 \wedge Z_1 = -Z_3)/\mathrm{P}(X_3 = 1)\\ &= 2p^2(1-p)/3p^2(1-p) = 2/3\end{aligned}$$

Exercise 4.3 (Random Walk). *Let V be a random variable and $V = 1$ with probability $1/2$ and $V = 1/2$ with probability $1/2$. Consider a random walk $X_t = \sum_{k=1}^{t} Z_k$ for $t = 1, 2, \ldots$, $X_0 = 0$, where $Z_1, Z_2, \ldots$ are i.i.d. with $P(Z_k = 1) = V$ and $P(Z_k = -1) = 1 - V$. Think as if one would toss a coin at an inception and on tails X_t would follow the ordinary random walk, while on heads it will deterministically increase. Calculate*

(a) $\mathrm{P}(X_t > 0)$
(b) $\mathrm{P}(X_t = 1)$
(c) $\mathrm{P}(Z_2 = 1|X_3 = 1)$

(a) $\mathrm{P}(X_t > 0) = \mathrm{P}(X_t > 0|V = 1)\mathrm{P}(V = 1) + \mathrm{P}(X_t > 0|V = 1/2)\mathrm{P}(V = 1/2) = 1/2 + \sum_{k=\lceil t/2+1\rceil}^{t} \binom{t}{k} \frac{1}{2}^{t+1}$, where $\lceil x \rceil$ means the largest integer not larger than x.
(b) For $t = 1$ $\mathrm{P}(X_t = 1) = 1/2 + (1/2) \cdot (1/2) = 3/4$, and for other odd t $\mathrm{P}(X_t = 1) = \binom{t}{\frac{1+t}{2}} \frac{1}{2}^{t+1}$
(c) Since $X_3 = 1$ can happen only if $V = 1/2$ then $\mathrm{P}(Z_2 = 1|X_3 = 1) = \mathrm{P}(Z_2 = 1|X_3 = 1 \wedge V = 1/2)\mathrm{P}(V = 1/2) = 1/3$.

Exercise 4.4 (Random Walk). *Consider an ordinary random walk* $X_t = \sum_{k=1}^{t} Z_k$ *for* $t = 1, 2, \ldots$; $X_0 = 0$, *where* $Z_1, Z_2, \ldots$ *are i.i.d. with* $\mathrm{P}(Z_i = 1) = p$ *and* $\mathrm{P}(Z_i = -1) = 1 - p$. *Let* $\tau = \min\{t : |X_t| > 1\}$ *be a random variable denoting the first time t when* $|X_t| > 1$. *Calculate* $\mathsf{E}\,\tau$.

It is easy to observe that $\mathrm{P}(\tau = 2k + 1) = 0$ for $k = 0, 1, \ldots$ and hence $X_\tau = 2$ or -2. One can then obtain

$$\begin{aligned} \mathrm{P}(\tau = 2) &= p^2 + (1 - p)^2 = q \\ \mathrm{P}(\tau = 4) &= \{1 - \mathrm{P}(\tau = 2)\}\mathrm{P}(\tau = 2) = (1 - q)q \end{aligned} \tag{4.1}$$

The first term in (4.1) corresponds to the probability that $\tau > 2$. The second term corresponds to the two consecutive up or down movements given that $X_2 = 0$. Similarly

$$\mathrm{P}(\tau = 6) = \{1 - \mathrm{P}(\tau = 2) - \mathrm{P}(\tau = 4)\}q = q(1 - q)^2$$

If $\mathrm{P}(\tau = 2k) = q(1 - q)^{k-1}$ then

$$\begin{aligned} \mathrm{P}(\tau = 2k + 2) &= \{1 - \mathrm{P}(\tau = 2) - \ldots - \mathrm{P}(\tau = 2k)\}q \\ &= \{1 - q - q(1 - q) - \ldots - q(1 - q)^{k-1}\}q \\ &= \left\{1 - q\frac{1 - (1 - q)^k}{q}\right\} q \\ &= q(1 - q)^k. \end{aligned}$$

Using induction $\mathrm{P}(\tau = 2k) = q(1 - q)^{k-1}$ for $k = 1, 2, \ldots$. Therefore τ has a geometric distribution with parameter $q = p^2 + (1 - p)^2$ and $\mathsf{E}\,\tau = 2\{p^2 + (1 - p)^2\}^{-1}$.

Exercise 4.5 (Random Walk). *Consider an ordinary random walk* $X_t = \sum_{k=1}^{t} Z_k$ *for* $t = 1, 2, \ldots,$ $X_0 = 0$, *where* $Z_1, Z_2, \ldots$ *are i.i.d. with* $\mathrm{P}(Z_i = 1) = p$ *and* $\mathrm{P}(Z_i = -1) = 1 - p$. *Consider also a process* $M_t = \max_{s \le t} X_s$. *Calculate*

(a) $\mathrm{P}(X_3 = M_3)$
(b) $\mathrm{P}(M_4 > M_3)$

(a) $\{X_3 = M_3\} = \{Z_1 = Z_2 = Z_3 = 1 \vee Z_1 = -1, Z_2 = Z_3 = 1 \vee Z_1 = Z_3 = 1, Z_2 = -1\}$ and hence $\mathrm{P}(X_3 = M_3) = p^3 + 2p(1-p)$.
(b) $\mathrm{P}(M_4 > M_3) = \mathrm{P}(X_3 = M_3)\mathrm{P}(Z_4 = 1) = p^4 + 2p^2(1-p)$

Exercise 4.6 (Random Walk). *Let* $X_t =, \sum_{k=1}^{t} Z_k$ *be a general random walk for* $t = 1, 2, \ldots,$ $X_0 = 0$, *and* $Z_1, Z_2, \ldots$ *are i.i.d. with* $\mathsf{Var}\, Z_i = 1$. *Calculate* $\mathsf{Cov}(X_s, X_t)$.

$$\mathsf{Cov}(X_s, X_t) = \mathsf{Cov}(\sum_{i=1}^{s} Z_i, \sum_{j=1}^{t} Z_j) = \sum_{i,j} \mathsf{Cov}(Z_i, Z_j) = \min(s,t)\, \mathsf{Var}\, Z_1 = \min(s,t)$$

Exercise 4.7 (Random Walk). *Let* $X_t = \sum_{k=1}^{t} Z_k$ *be a general random walk for* $t = 1, 2, \ldots,$ $X_0 = 0$, *and* $Z_1, Z_2, \ldots$ *are i.i.d. and symmetric random variables. Show that*

$$\mathrm{P}(\max_{i \le t} |X_i| > a) \le 2\mathrm{P}(|X_t| > a).$$

Denote an event that the level t is breached for the first time in the i-th step by $A_i = \{|X_j| \le t \text{ for } j = 1, 2, \ldots, i-1, |X_i| > t\}$. One may show that

$$A_i \subset (A_i \cap |X_t| > a) \cup (A_i \cap |2X_i - X_t| > a)$$

because given that $|X_i| > a$ then $|2X_i - X_t| > a$ or $|X_t| > a$ since

$$2a < |X_t + 2X_i - X_t| < |X_t| + |2X_i - X_t|.$$

Note that $X_t = X_i + X_t - X_i$ and $2X_i - X_t = X_i - (X_t - X_i)$ have the same distribution because of the independence and symmetry of Z_i. Therefore

$$\mathrm{P}(A_i) \le \mathrm{P}(A_i \cap |X_t| > a) + \mathrm{P}(A_i \cap |2X_i - X_t| > a) = 2\mathrm{P}(A_i \cap |X_t| > a)$$

and

$$\begin{aligned} \mathrm{P}(\max_{i \le t} X_i > a) = \sum_{i=1}^{t} \mathrm{P}(A_i) &\le 2 \sum_{i=1}^{t} \mathrm{P}(A_i \cap |X_t| > a) \\ &= 2\mathrm{P}(\max_{i \le n} X_i > a, |X_t| > a) \\ &\le 2\mathrm{P}(|X_t| > a) \end{aligned}$$

Exercise 4.8 (Binomial Process). *Consider a binomial process* $X_t = \sum_{k=1}^{t} Z_k$ *for* $t = 1, 2, \ldots$, $X_0 = 0$, *with state dependent increments. Let* $\mathrm{P}(Z_t = 1) = 1/(2^{|X_{t-1}|+1})$ *if* $X_{t-1} \geq 0$ *and* $\mathrm{P}(Z_t = 1) = 1 - 1/(2^{|X_{t-1}|+1})$ *otherwise. To complete the setting* $\mathrm{P}(Z_t = -1) = 1 - \mathrm{P}(Z_t = 1)$. *Calculate the distribution of* X_t *for the first five steps.*

As

$$\mathrm{P}(Z_t = -1) = \begin{cases} 1 - 1/(2^{|X_{t-1}|+1}) & \text{if } X_{t-1} \geq 0 \\ 1/(2^{|X_{t-1}|+1}) & \text{if } X_{t-1} < 0 \end{cases}$$

is equivalent to:

$$\mathrm{P}(Z_t = 1) = \begin{cases} 1/(2^{|X_{t-1}|+1}) & \text{if } X_{t-1} \geq 0 \\ 1 - 1/(2^{|X_{t-1}|+1}), & \text{if } X_{t-1} < 0 \end{cases}$$

The table of states probabilities must be symmetric. Therefore we only have to consider cases where $X_t \geq 0$.

When $t = 1$,

$$\mathrm{P}(Z_t = -1) = \mathrm{P}(Z_t = 1) = 1/2$$
$$\mathrm{P}(X_t = 1) = \mathrm{P}(X_t = -1) = 1/2$$

When $t = 2$; $X_t = 1$

$$\mathrm{P}(Z_t = 1) = 1/4$$
$$\mathrm{P}(Z_t = -1) = 3/4$$

According to the symmetry, we have:

$$\mathrm{P}(X_t = 2) = \mathrm{P}(X_t = -2) = 1/2 \cdot 1/4 = 1/8$$
$$\mathrm{P}(X_t = 0) = 1/2 \cdot 3/4 \cdot 2 = 3/4$$

Calculations for $t = 3, \ldots$ are quite similar and are not covered here.

The distribution of the first five steps could conveniently be illustrated by the following table of states probabilities. Note that with this construction of the probabilities the process tends to level 0 (Table 4.2).

Exercise 4.9 (Geometric Binomial Process). *Suppose* X_t *is a geometric binomial process with* $X_0 = 1$. *Further the return* $R_t = X_t / X_{t-1}$ *is identically and independently log-normal distributed:* $R_t \sim$ *lognormal*$(0, 1)$. *Calculate the expected value* $\mathsf{E}[X_6 | X_4 = 1, X_3 = 2]$.

Table 4.2 The distribution of X_t for the first five steps. SFS5step

X_t	Probabilities					
5						1/32,768
4					1/1,024	
3				1/64		587/32,768
2			1/8		139/1,024	
1		1/2		31/64		3,949/8,192
0	1		3/4		186/256	
−1		1/2		31/64		3,949/8,192
−2			1/8		139/1,024	
−3				1/64		587/32,768
−4					1/1,024	
−5						1/32,768
t	0	1	2	3	4	5

As the return $R_t = X_t/X_{t-1}$ is i.i.d., $X_6 = X_4 R_5 R_6 = R_5 R_6$, we have $\mathsf{E}[X_6|X_4 = 1, X_3 = 2] = \mathsf{E}[R_5]\,\mathsf{E}[R_6]$. From the property of standard lognormal distribution, $\mathsf{E}[R_t] = \exp(0 + 0.5 \cdot 1) = \exp(1/2)$, so $\mathsf{E}[X_6] = \exp(1/2) = \exp(1/2) \cdot \exp(1/2) = e$

Chapter 5
Stochastic Integrals and Differential Equations

Intégrals Stochastique et Équations Différentielle
Prudence est mère de sûreté
Discretion is the better part of valour

In the preceding chapter we discussed stochastic processes in discrete time. This chapter is devoted to stochastic processes in continuous time. An important continuous time process is the standard Wiener process $\{W_t;\ t \geq 0\}$. For this process it holds for all $0 \leq s \leq t$:

$$\begin{aligned} \mathsf{E}[W_t] &= 0,\ \mathsf{Var}(W_t) = t \\ \mathsf{Cov}(W_t, W_s) &= \mathsf{Cov}(W_t - W_s + W_s, W_s) \\ &= \mathsf{Cov}(W_t - W_s, W_s) + \mathsf{Cov}(W_s, W_s) \\ &= 0 + \mathsf{Var}(W_s) = s \end{aligned}$$

In the context of this chapter we also consider the Itô process $\{X_t;\ t \geq 0\}$:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \tag{5.1}$$

The time and state dependent terms μ and σ represent the drift rate and the variance respectively. A precise definition of a solution to (5.1) is a stochastic process fulfilling the integral equation:

$$X_t - X_0 = \int_0^t \mu(X_s, s)ds + \int_0^t \sigma(X_s, s)dW_s \tag{5.2}$$

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_5,

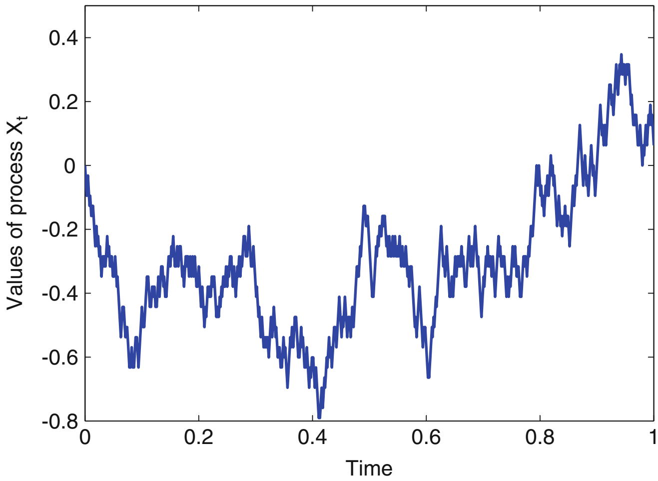


Fig. 5.1 Graphic representation of a standard Wiener process X_t on 1,000 equidistant points in interval [0, 1]. SFSwiener1

Exercise 5.1 (Wiener Process). *Let W_t be a standard Wiener process. Show that the following processes are also standard Wiener processes:*

(a) $X_t = c^{-\frac{1}{2}} W_{ct}$ *for* $c > 0$
(b) $Y_t = W_{T+t} - W_T$ *for* $T > 0$
(c) $V_t = \begin{cases} W_t & \text{if } t \leq T \\ 2W_T - W_t & \text{if } t > T \end{cases}$

It is easy to check that all processes start at 0, have a zero mean and independent increments since W_t has independent increments. One has to additionally check the variance of the increments for $t > s \geq 0$.

(a) $\mathsf{Var}(X_t - X_s) = \mathsf{Var}(c^{-\frac{1}{2}} W_{ct} - c^{-\frac{1}{2}} W_{cs}) = c^{-1}(ct - cs) = t - s$ (b) $\mathsf{Var}(Y_t - Y_s) = \mathsf{Var}(W_{T+t} - W_T - W_{T+s} + W_T) = \mathsf{Var}(W_{T+t} - W_{T+s}) = t - s$ (c) For $s < t < T$ and $T < s < t$ one directly obtains the increments of W_t. For $s < T < t$ one has:

$$\mathsf{Var}(V_t - V_s) = \mathsf{Var}(2W_T - W_t - W_s) = \mathsf{Var}(W_T - W_t) + \mathsf{Var}(W_T - W_s) = t - T + T - s = t - s.$$

The additivity of variance follows from the independent increments of W_t.

Corresponding to (a), Fig. 5.1 gives the plot of X_t on 1,000 equisidtant points in intervall [0, 1] with $c = 2$. Similar plots for (b) and (c) are omitted here, and detailed codes could be found in SFSwiener2 and SFSwiener3.

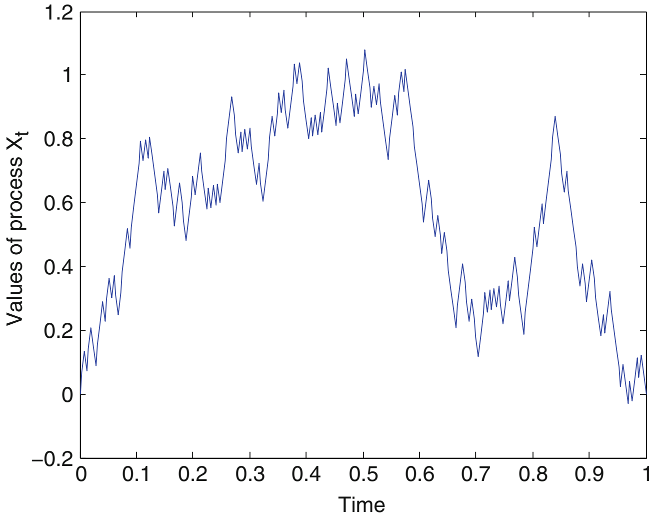


Fig. 5.2 A Brownian bridge. SFSbb

Exercise 5.2 (Covariance). *Calculate* $\mathsf{Cov}(2W_t, 3W_s - 4W_t)$ *and* $\mathsf{Cov}(2W_s, 3W_s - 4W_t)$ *for* $0 \le s \le t$.

$$\mathsf{Cov}(2W_t, 3W_s - 4W_t) = \mathsf{Cov}(2W_t, 3W_s) - \mathsf{Cov}(2W_t, 4W_t) = 6s - 8t.$$

$$\mathsf{Cov}(2W_s, 3W_s - 4W_t) = \mathsf{Cov}(2W_s, 3W_s) - \mathsf{Cov}(2W_s, 4W_t) = 6s - 8s = -2s.$$

Exercise 5.3 (Brownian Bridge). *Let* W_t *be a standard Wiener process. The process* $U_t = W_t - tW_1$ *for* $t \in [0, 1]$ *is called Brownian bridge. Calculate its covariance function. What is the distribution of* U_t.

$$\mathsf{Cov}(U_t, U_s) = \mathsf{Cov}(W_t - tW_1, W_s - sW_1) = \mathsf{Cov}(W_t, W_s) + ts\,\mathsf{Cov}(W_1, W_1) - t\,\mathsf{Cov}(W_1, W_s) - s\,\mathsf{Cov}(W_t, W_1) = \min(t, s) - ts.$$

The distribution of U_t is normal with mean 0 and variance: $\mathsf{Var}(W_t - tW_1) = \mathsf{Var}(W_t) + 2\ \mathsf{Cov}(W_t, -tW_1) + \mathsf{Var}(-tW_1) = t - 2t^2 + t^2 = t(1-t)$

Figure 5.2 displays one example of a Brownian bridge.

Exercise 5.4 (Reflection Property). *Using the reflection property (see Exercise 4.6), i.e.* $\mathrm{P}(\sup_{s \le t} W_s > x) = 2\mathrm{P}(W_t > x)$ *for* $x \ge 0$, *calculate the density of* $\sup_{s \le t} W_s$.

$$\mathrm{P}(\sup_{s \le t} W_s \le x) = 1 - 2\mathrm{P}(W_t > x) = 2\mathrm{P}(W_t \le x) - 1 = 2\Phi(x/\sqrt{t}) - 1.$$

This result implies in particular that the Wiener process has both positive and negative values on interval $[0, t]$ for each t.

The density of $M_t = \sup_{s \leq t} W_s$ is then given by:
$f_{\sup W}(x) = 2\varphi(x/\sqrt{t})/\sqrt{t} = \sqrt{2/(\pi t)}\exp\{-x^2/(2t)\}$ for $x \geq 0$ and $f_{\sup W}(x) = 0$ otherwise.

Exercise 5.5 (Integration). *Calculate* $\mathsf{E}\left(\int_0^{2\pi} W_s dW_s\right)$

According to the rule of integration by parts, we have: $\int_0^{2\pi} W_s dW_s = \frac{1}{2}(W_{2\pi}^2 - 2\pi)$. Together with $\mathsf{Var}(W_{2\pi}) = \mathsf{E}\, W_{2\pi}^2 - \mathsf{E}^2\, W_{2\pi}$, we get:

$$\mathsf{E}\,\frac{1}{2}(W_{2\pi}^2 - 2\pi) = \frac{1}{2}\ \mathsf{Var}(W_{2\pi}) - \pi = 0$$

Exercise 5.6 (Brownian Motion). *Find the dynamics of* $Y_t = \sin(W_t)$ *for a Brownian motion* W_t*.*

According to Itô's lemma:

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

together with $X_t = W_t,\ g(X_t) = \sin(W_t),\ \mu = 0, \sigma = 1$, we have $dY_t = \cos(X_t)dW_t - 0.5\sin(W_t)dt$

Exercise 5.7 (Dynamics). *Consider the process* $dS_t = \mu dt + \sigma dW_t$*. Find the dynamics of the process* $Y_t = g(S_t)$*, where* $g(S_t, t) = 2 + t + \exp(S_t)$*.*

According to Itô's lemma:

$$\begin{aligned} dY_t &= \frac{\partial g}{\partial s}dS_t + \frac{\partial g}{\partial t}dt + 0.5\frac{\partial^2 g}{\partial s^2}(dS_t)^2 + 0 \\ &= \exp(S_t)(\mu dt + \sigma dW_t) + dt + 0.5\sigma^2\exp(S_t)dt \\ &= \{1 + (\mu + 0.5\sigma^2)\exp(S_t)\}dt + \sigma\exp(S_t)dW_t \end{aligned}$$

Exercise 5.8 (Brownian Motion). *Derive* $\int_0^t W_s^2 dW_s$*, where* W_t *is a Brownian motion.*

Choose $Y_t = \frac{1}{3}W_t^3$. According to Itô's lemma:

$$dY_t = W_t^2 dW_t + W_t dt.$$

Thus $Y_t = \int_0^t W_s^2 dW_s + \int_0^t W_s ds$ and hence

$$\int_0^t W^2 dW_s = \frac{1}{3}W_t^3 - \int_0^t W_s ds.$$

Exercise 5.9 (Wiener Process). *Let W_t be a standard Wiener process. Compute then $E[W_t^4]$.*

Let $X_t = W_t^4$, $g(X) = X^4$, $g'(X) = 4X^3$, $g''(X) = 12X^2$. By applying Itô's lemma we get:

$$\begin{aligned} dX_t &= (4W_t^3 \cdot 0 + \frac{1}{2} \cdot 12W_t^2 \cdot 1)dt + 4W_t^3 dW_t \\ &= 4W_t^3 dW_t + 6W_t^2 dt \end{aligned}$$

Integrating both parts, we get:

$$\begin{aligned} X_t - X_0 &= 4\int_0^t W_s^3 dW_s + 6\int_0^t W_s^2 ds \\ X_t &= 4\int_0^t W_s^3 dW_s + 6\int_0^t W_s^2 ds \end{aligned}$$

Computing the expectation leads to:

$$\begin{aligned} \mathsf{E}(X_t) = \mathsf{E}(W_t^4) &= 4\ \mathsf{E}\left(\int_0^t W_s^3 dW_s\right) + 6\,\mathsf{E}\left(\int_0^t W_s^2 ds\right) \\ &= 6\int_0^t \mathsf{E}(W_s^2) ds = 6\int_0^t s ds = 6\left(\frac{t^2}{2}\right) = 3t^2 \end{aligned}$$

Exercise 5.10 (Differential Equation). *If $g = g(y)$ is a function of y, and suppose $y = f(w)$ is the solution of the following ordinary differential equation:*

$$dy = g(y)dw$$

Show that $X_t = f(W_t)$ is a solution of the stochastic differential equation:

$$dX_t = \frac{1}{2} g(X_t) g'(X_t) dt + g(X_t) dW_t$$

If $X_t = f(W_t)$, then by applying Itô's lemma we obtain the following result:

$$\begin{aligned} dX_t &= df(W_t) \\ &= \left(\frac{\partial f}{\partial W_t} \cdot 0 + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2} \cdot 1\right) dt + \frac{\partial f}{\partial W_t} \cdot 1 dW_t \\ &= \frac{\partial f}{\partial W_t} \cdot dW_t + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2} dt \\ &= f'(W_t) dW_t + \frac{1}{2} f''(W_t) dt. \end{aligned}$$

From the standard calculus

$$f'(w) = \frac{dy}{dw} = g(y) \text{ for } y = f(w),$$
$$f'(w) = g\{f(w)\},$$
$$f''(w) = g'\{f(w)\} \cdot f'(w)$$

and by substituting we get:

$$\begin{aligned} dX_t &= f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt \\ &= g\{f(W_t)\}dW_t + \frac{1}{2}f''(W_t)dt \\ &= g\{f(W_t)\}dW_t + \frac{1}{2}g'\{f(W_t)\}g\{f(W_t)\}dt \\ &= g(X_t)dW_t + \frac{1}{2}g'(X_t)g(X_t)dt \end{aligned}$$

Exercise 5.11 (Stochastic Differential Equation). *Apply the previous result to solve the following SDE's*

(a) $dX_t = \sqrt{X_t}dW_t + \frac{1}{4}dt$
(b) $dX_t = X_t^2 dW_t + X_t^3 dt$
(c) $dX_t = \cos^2 X_t dW_t - \frac{1}{2}(\sin 2X_t)\cos^2 X_t dt$

(a) One may easily check that the stochastic differential equation

$$dX_t = \sqrt{X_t}dW_t + \frac{1}{4}dt$$

is of the form discussed in Exercise 5.10

$$dX_t = g(X_t)dW_t + \frac{1}{2}g'(X_t)g(X_t)dt.$$

By comparing the term of dW_t one obtains:

$$g(X_t) = \sqrt{X_t}$$
$$\frac{1}{2}g'(X_t)g(X_t) = \frac{1}{2} \cdot \frac{1}{2\sqrt{X_t}} \cdot \sqrt{X_t} = \frac{1}{4}$$

In the next step one finds the function $X_t = f(W_t)$; this solves the ordinary differential equation

$$
\begin{aligned}
f'(w) &= g\{f(w)\} = \sqrt{f(w)} \\
f'(w) &= \frac{df}{dw} \\
\frac{df}{dw} &= \sqrt{f(w)} \\
\frac{df}{\sqrt{f}} &= dw
\end{aligned}
$$

Integrating both parts, it results in:

$$
\begin{aligned}
2\sqrt{f(w)} &= w + C \\
\sqrt{f(w)} &= \frac{w + C}{2} \\
f(w) &= \left(\frac{w + C}{2}\right)^2 \\
X_t &= \left(\frac{W_t + C}{2}\right)^2 .
\end{aligned}
$$

For $t = 0$:

$$
\begin{aligned}
X_0 &= \left(\frac{W_0 + C}{2}\right)^2; \;\; W_0 = 0 \\
X_0 &= \left(\frac{C}{2}\right)^2 \\
C &= 2\sqrt{X_0} \;\text{ with }\; X_0 \geq 0.
\end{aligned}
$$

Therefore the solution is:

$$
X_t = \left(\frac{W_t + 2\sqrt{X_0}}{2}\right)^2
$$

(b) Here the function $g(x)$ has the form

$$
\begin{aligned}
g(X_t) &= X_t^2 \\
g'(X_t) &= 2X_t \\
\frac{1}{2}g'(X_t) \cdot g(X_t) &= \frac{1}{2}X_t^2 \cdot 2X_t \;\; = \;\; X_t^3 .
\end{aligned}
$$

To find f one has

$$\begin{aligned}
f'(w) &= g\{f(w)\} = \{f(w)\}^2 \\
\frac{df}{dw} &= f^2(w) \\
\frac{df}{f^2(w)} &= dw \\
-d\left(\frac{1}{f(w)}\right) &= dw
\end{aligned}$$

Integrating both parts gives:

$$\begin{aligned}
-1/f(w) &= w + C \\
f(w) &= -1/(w + C) \\
X_t &= -1/(W_t + C).
\end{aligned}$$

In order to determine the constant C, we use the initial condition

$$\begin{aligned}
X_0 &= -1/(W_0 + C),\ W_0 = 0 \\
C &= -1/X_0.
\end{aligned}$$

The solution is then:

$$X_t = -1/(W_t - 1/X_0)$$

(c) In this example function g has the form

$$\begin{aligned}
g(X_t) &= \cos^2 X_t \\
g'(X_t) &= -2\cos X_t \sin X_t \quad = \quad -\sin 2X_t \\
\frac{1}{2}g'(X_t)\cdot g(X_t) &= -\frac{1}{2}\sin 2X_t \cdot \cos^2 X_t.
\end{aligned}$$

We solve the ordinary differential equation

$$\begin{aligned}
f'(w) &= g\{f(w)\} = \cos^2 f(w) \\
\frac{df}{dw} &= \cos^2 f(w) \\
\frac{df}{\cos^2 f(w)} &= dw \\
d\tan f(w) &= dw
\end{aligned}$$

Integrating both parts we obtain:

$$\tan f(W_t) = W_t + C$$
$$X_t = f(W_t) = \arctan(W_t + C)$$

The initial condition gives:

$$X_0 = \arctan(W_0 + C), \ W_0 = 0$$
$$X_0 = \arctan C$$
$$C = \tan X_0$$

The solution is then:

$$X_t = \arctan(W_t + \tan X_0)$$

Exercise 5.12 (Martingale). *Let B_t be an Itô process and $M_t = f(t, B_t)$ where $f(t,x) = \exp(x)\cos(x + at)$. Use Itô's lemma to determine a constant a so that*

$$M_t = \exp(B_t)\cos(B_t + at)$$

is a martingale.

Hint: to show that M_t is a martingale, one has to show that M_t is of the form

$$M_t = \int_0^t g(s, B_s) dB_s$$

and that

$$\mathsf{E}\left[\left\{\int_0^t g(s, B_s) dB_s\right\}^2\right] < \infty$$

See the Novikov condition in Franke et al. (2011).

We apply Itô's lemma to derive at first dM_t.

If $f(t,x) = \exp(x)\cos(x + at)$, then

$$\frac{\partial f(t,x)}{\partial t} = -\exp(x) \cdot a \cdot \sin(x + at)$$
$$\frac{\partial f(t,x)}{\partial x} = \exp(x)\cos(x + at) - \exp(x)\sin(x + at)$$
$$\begin{aligned}\frac{\partial^2 f(t,x)}{\partial x^2} &= \exp(x)\cos(x + at) - \exp(x)\sin(x + at) \\ &\quad -\{\exp(x)\sin(x + at) + \exp(x)\cos(x + at)\} \\ &= \exp(x)\{\cos(x + at) - \sin(x + at) - \sin(x + at) - \cos(x + at)\} \\ &= -2\exp(x)\sin(x + at)\end{aligned}$$

and

$$\begin{aligned} dM_t &= \frac{\partial f(t, B_t)}{\partial t} \\ &= \frac{\partial f(t,x)}{\partial t}|_{x=B_t} dt + \frac{\partial f(t,x)}{\partial x}|_{x=B_t} dB_t + \frac{1}{2}\frac{\partial^2 f(t,x)}{\partial x^2}|_{x=B_t} dt \\ &= \exp(B_t)\{\cos(B_t + at) - \sin(B_t + at)\} dB_t \\ &\quad + \exp(B_t)\{-a \sin(B_t + at) - \sin(B_t + at)\} dt \\ M_t - M_0 &= \int_0^t \exp(B_s)\{\cos(B_s + as) - \sin(B_s + as)\} dB_s \\ &\quad + \int_0^t \exp(B_s)\{-a \sin(B_s + as) - \sin(B_s + as)\} ds. \end{aligned}$$

Additionally

$$\begin{aligned} M_0 &= \exp(B_0)\cos(B_0 + 0) \\ &= \exp(0) \cdot \cos 0 = 1 \end{aligned}$$

M_t is a martingale, if a satisfies:

$$\int_0^t \exp(B_s)\{-a \sin(B_s + as) - \sin(B_s + as)\} ds = 0.$$

such that

$$M_t = 1 + \int_0^t \exp(B_s)\{\cos(B_s - s) - \sin(B_s - s)\} dB_s$$

is a stochastic integral.

To show that M_t is a martingale, we also have to show that:

$$\mathsf{E}\left(\left[\int_0^t \exp(B_s)\{\cos(B_s + s) - \sin(B_s - s)\} dB_s\right]^2\right) < \infty$$

If we write $\exp(B_s)\{\cos(B_s + s) - \sin(B_s - s)\}$ as $h(s)$, together with

$$\mathsf{E}\left\{\int_0^t h(s) dB_s\right\}^2 = \int_0^t \{\mathsf{E}(h^2(s)\} ds,$$

one obtains:

$$\mathsf{E}\left(\left[\int_0^t \exp(B_s)\left\{\cos(B_s+s)-\sin(B_s-s)\right\}dB_s\right]^2\right)$$

$$= \int_0^t \mathsf{E}\left[\exp(B_s)\underbrace{\{\cos(B_s-s)-\sin(B_s-s)\}}_{\leq 2}\right]^2 ds$$

$$\leq \int_0^t \mathsf{E}\left\{\exp(2B_s)\cdot 2^2\right\} ds$$

$$= 4\int_0^t \mathsf{E}\left\{\exp(2B_s)\right\} ds$$

$$= 4\int_0^t \exp(2s)ds < \infty$$

Here we use the formula for the expectation of a geometric Brownian motion $\mathsf{E}\{\exp(2B_s)\} = \exp(2\mu s + \sigma^2 s)$, where μ and σ are the mean and standard deviation of B_s. In our case $\mu = 0$ and $\sigma = 1$, due to the fact that B_s is a standard Wiener process.

The set of a such that M_t is a martingale is:

$$\{a| \int_0^t \exp(B_s)\{-a\sin(B_s+as)-\sin(B_s+as)\}ds = 0\}$$

Exercise 5.13 (Product Rule and Integration by Parts). *If X_t and Y_t are two one-dimensional Itô processes with:*

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s$$

$$Y_t = Y_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$

Then we have:

$$X_t \cdot Y_t = X_0\cdot Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t dX_t dY_t$$

$$= X_0\cdot Y_0 + \int_0^t (X_s\mu_s + Y_sK_s + H_s\sigma_s)ds + \int_s^t (H_s\sigma_s + Y_sH_s)dW_s$$

Differentiating both sides leads to:

$$d(X_tY_t) = X_tdY_t + Y_tdX_t + dX_tdY_t.$$

This expression represents another form of the product rule, in differential form. The differential form of the two Itô processes is

$$\begin{aligned} dX_t &= K_t dt + H_t dW_t \\ dY_t &= \mu_t dt + \sigma_t dW_t \\ d(X_t + Y_t)^2 &= d(X_t^2 + 2X_t Y_t + Y_t^2) \\ &= dX_t^2 + 2d(X_t Y_t) + dY_t^2 \end{aligned}$$

In order to derive the expressions for dX_t^2 and dY_t^2 we apply Itô's lemma. Let

$$g(X) = X^2 \qquad g'(X) = 2X \qquad g''(X) = 2,$$

then

$$\begin{aligned} dX_t^2 &= dg(X_t) \\ &= (2X_t K_t + \frac{1}{2} 2H_t^2)dt + 2X_t H_t dW_t \\ &= (2X_t K_t + H_t^2)dt + 2X_t H_t dW_t \end{aligned}$$

Analogously, we obtain for dY_t^2

$$dY_t^2 = dg(Y_t) = (2Y_t \mu_t + \sigma_t^2)dt + 2Y_t \sigma_t dW_t$$

Let

$$\begin{aligned} Z_t &= Y_t + X_t \\ &= X_0 + Y_0 + \int_0^t (K_s + \mu_s)ds + \int_0^t (H_s + \sigma_s)dW_s \\ dZ_t &= (K_t + \mu_t)dt + (H_t + \sigma_t)dW_t \end{aligned}$$

then, by applying the same principle as above, we get:

$$\begin{aligned} dZ_t^2 &= \left\{2Z_t(K_t + \mu_t) + (H_t + \sigma_t)^2\right\} dt + 2Z_t(H_t + \sigma_t)dW_t \\ &= \left\{2(X_t + Y_t)(K_t + \mu_t) + (H_t + \sigma_t)^2\right\} dt + 2(Y_t + X_t)(H_t + \sigma_t)dW_t \\ 2d(X_t Y_t) &= dZ_t^2 - dX_t^2 - dY_t^2 \\ &= \left\{2(X_t + Y_t)(K_t + \mu_t) + (H_t + \sigma_t)^2\right\} dt + 2(Y_t + X_t)(H_t + \sigma_t)dW_t \\ &\quad - \left\{(2X_t K_t + H_t^2)dt + 2X_t H_t dW_t\right\} - \left\{(2Y_t \mu_t + \sigma_t^2)dt + 2Y_t \sigma_t dW_t\right\} \\ &= \{2(X_t K_t + X_t \mu_t + Y_t K_t + Y_t \mu_t - X_t K_t - Y_t \mu_t) \\ &\quad + (H_t^2 + \sigma_t^2 + 2H_t \sigma_t - H_t^2 - \sigma_t^2)\}dt \\ &\quad + 2(Y_t H_t + Y_t \sigma_t + X_t H_t + X_t \sigma_t - X_t H_t - Y_t \sigma_t)dW_t \end{aligned}$$

$$
\begin{aligned}
&= 2(X_t\mu_t + Y_t K_t + H_t\sigma_t)dt + 2(Y_t H_t + X_t\sigma_t)dW_t \\
d(X_t Y_t) &= (X_t\mu_t + Y_t K_t + H_t\sigma_t)dt + (Y_t H_t + X_t\sigma_t)dW_t \\
&= X_t\mu_t dt + Y_t K_t dt + H_t\sigma_t dt + Y_t H_t dW_t + X_t\sigma_t dW_t \\
&= X_t(\mu_t dt + \sigma_t dW_t) + Y_t(K_t dt + H_t dW_t) + H_t\sigma_t dt \\
&= X_t dY_t + Y_t dX_t + H_t\sigma_t dt
\end{aligned}
$$

In order to prove the equality above, we still have to show that

$$
\begin{aligned}
dX_t dY_t &= H_t\sigma_t dt \\
\text{Hence,}\quad dX_t \cdot dY_t &= (K_t dt + H_t dW_t)(\mu_t dt + \sigma_t dW_t) \\
&= K_t\mu_t (dt)^2 + K_t\sigma_t dt\, dW_t + H_t\mu_t dW_t dt + H_t\sigma_t (dW_t)^2
\end{aligned}
$$

The terms $(dt)^2$ and $dt\, dW_t$ are $o(dt)$, together with $(dW_t)^2 = dt$ it follows then:

$$
\begin{aligned}
dX_t dY_t &= H_t\sigma_t (dW_t)^2 \\
&= H_t\sigma_t dt
\end{aligned}
$$

We have shown that:

$$
d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t
$$

by integrating back, we obtain:

$$
\begin{aligned}
X_t Y_t &= Y_0 X_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \int_0^t dX_s dY_s \\
&= Y_0 X_0 + \int_0^t X_s(\mu_s ds + \sigma_s dW_s) + \int_0^t Y_s(K_s ds + H_s dW_s) + \int_0^t H_s\sigma_s ds \\
&= Y_0 X_0 + \int_0^t (X_s\mu_s + Y_s K_s + H_s\sigma_s)ds + \int_0^t (X_s\sigma_s + Y_s H_s)dW_s
\end{aligned}
$$

Exercise 5.14 (Ornstein-Uhlenbeck process). *Prove that the following process:*

$$
S_t = \exp(-\mu t)S_0 + \theta\{1 - \exp(-\mu t)\} + \exp(-\mu t)\int_0^t \sigma \exp(\mu s)dW_s \tag{5.3}
$$

is a solution of the Ornstein-Uhlenbeck SDE:

$$
dS_t = -\mu(S_t - \theta)dt + \sigma dW_t \tag{5.4}
$$

where θ, μ and σ are parameters and W_t is a standard Wiener process.

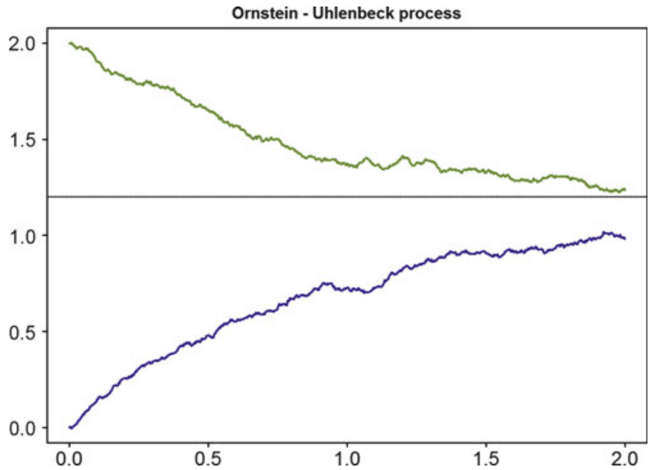


Fig. 5.3 Graphic representation of an Ornstein-Uhlenbeck process with different initial values. SFSornstein

Starting from (5.3) we have

$$\begin{aligned} S_t &= \exp(-\mu t)\left(S_0 + \int_0^t \exp(\mu s)\sigma dW_s\right) + \theta\{1 - \exp(-\mu t)\} \\ &= \exp(-\mu t)(S_0 + X_t - \theta) + \theta \end{aligned}$$

with

$$\begin{aligned} X_t &= \int_0^t \exp(\mu s)\sigma dW_s \\ dX_t &= \exp(\mu t)\sigma dW_t. \end{aligned}$$

By differentiating both sides we obtain:

$$\begin{aligned} dS_t &= d\{\exp(-\mu t)(S_0 + X_t) - \theta\} \\ &= -\mu \exp(-\mu t)(S_0 + X_t - \theta)dt + \exp(-\mu t)dX_t \\ &= -\mu \underbrace{\exp(-\mu t)(S_0 + X_t - \theta)}_{S_t - \theta} dt + \exp(-\mu t)\exp(\mu t)\sigma dW_t \\ &= -\mu(S_t - \theta)dt + \sigma dW_t. \end{aligned}$$

To see the intuition of this process, W.L.G. we choose $\mu = 1$, $\theta = 1.2$ and $\sigma = 0.3$. Figure 5.3 displays the plot. Dark green and blue lines correspond to different initial values 2 and 0.

Exercise 5.15 (Stochastic Differential Equation). *Apply the result from Exercise 5.14 to solve the following SDE:*

$$dX_t = aX_t(\theta - \log X_t)dt + \sigma X_t dW_t \tag{5.5}$$

Dividing both sides of (5.5) by X_t we obtain:

$$\frac{dX_t}{X_t} = a(\theta - \log X_t)dt + \sigma dW_t$$

Define now:

$$Y_t \overset{\text{def}}{=} g(X_t) = \log X_t$$

and observe that:

$$g'(X_t) = 1/X_t \quad g''(X_t) = -1/X_t^2$$

Comparing coefficients with an Itô process, we get:

$$\begin{aligned}\mu(X_t) &= aX_t(\theta - \log X_t)\\ \sigma(X_t) &= \sigma \cdot X_t\end{aligned}$$

By applying Itô's lemma, we obtain:

$$\begin{aligned}dY_t &= \left\{\frac{\partial g}{\partial X}(X_t)\cdot\mu(X_t) + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}(X_t)\cdot\sigma^2(X_t)\right\}dt + \frac{\partial g}{\partial X}(X_t)\cdot\sigma(X_t)dW_t\\ &= \left\{\frac{1}{X_t}aX_t(\theta - \log X_t) + \frac{1}{2}(-\frac{1}{X_t^2})\sigma^2 X_t^2\right\}dt + \frac{1}{X_t}\sigma X_t dW_t\\ &= \left\{a(\theta - \log X_t) - \frac{1}{2}\sigma^2\right\}dt + \sigma dW_t\\ &= a\left(\theta - \log X_t - \frac{\sigma^2}{2a}\right)dt + \sigma dW_t\\ &= a\left(\theta - \frac{\sigma^2}{2a} - Y_t\right)dt + \sigma dW_t\end{aligned}$$

Hence, Y_t follows the path of an Ornstein-Uhlenbeck process (5.4).

$$\begin{aligned}dY_t &= a(\theta - \frac{\sigma^2}{2a} - Y_t)dt + \sigma dW_t\\ &= -a\left\{Y_t - (\theta - \frac{\sigma^2}{2a})\right\}dt + \sigma dW_t\end{aligned}$$

The solution of the SDE in Y_t is then:

$$\begin{aligned}
\exp(at)Y_t &= Y_0 + \int_0^t \exp(as)a\left(\theta - \frac{\sigma^2}{2a}\right)ds + \int_0^t \sigma\exp(\theta s)dW_s \\
&= Y_0 + \left(\theta - \frac{\sigma^2}{2a}\right)\exp(as)|_0^t + \int_0^t \sigma\exp(as)dW_s \\
&= Y_0 + \left(\theta - \frac{\sigma^2}{2a}\right)\{\exp(at) - 1\} + \int_0^t \sigma\exp(as)dW_s \\
Y_t &= \exp(-at)Y_0 + \left(\theta - \frac{\sigma^2}{2a}\right)\{1 - \exp(-at)\} + \exp(-at)\int_0^t \sigma\exp(as)dW_s
\end{aligned}$$

thus for $Y_t = \log X_t,\ Y_0 = \log X_0$ it follows that:

$$\begin{aligned}
X_t = \exp\bigg[&\exp(-at)\log X_0 + \left(\theta - \frac{\sigma^2}{2a}\right)\{1 - \exp(-at)\} \\
&+ \exp(-at)\int_0^t \sigma\exp(as)dW_s\bigg]
\end{aligned}$$

Chapter 6
Black-Scholes Option Pricing Model

Modelul Black-Scholes de Evaluare a Optiunilor
Ulciorul nu merge de multe ori la apă.
The pitcher goes so often to the well that it comes home broken at last.

The Black-Scholes formula is one of the most recognizable formulae in quantitative finance. The formula for the price $C(S, \tau)$ of a European call option is given by:

$$C(S, \tau) = \exp\{(b - r)\tau\} S\Phi(y + \sigma\sqrt{\tau}) - \exp(-r\tau) K\Phi(y), \tag{6.1}$$

where we use y as an abbreviation for

$$y = \frac{\log(S/K) + \left\{b - \sigma^2/2\right\}\tau}{\sigma\sqrt{\tau}} \tag{6.2}$$

and $b - r$ denotes the cost of carry b subtracted by the interest rate r.

The corresponding Black–Scholes formula for the price $P(S, \tau)$ of a European put option can be found, for example, by using the put–call parity as in Franke et al. (2011):

$$P(S, \tau) = C(S, \tau) - S\exp\{(b - r)\tau\} + K\exp(-r\tau).$$

From this and Eq. (6.1) we obtain

$$P(S, \tau) = \exp(-r\tau) K\Phi(-y) - \exp\{(b - r)\tau\} S\Phi(-y - \sigma\sqrt{\tau}). \tag{6.3}$$

The stop–loss strategy is a strategy to decrease the risk associated with a long call as an expensive hedging strategy, i.e. the bank selling the option takes an uncovered position as long as the stock price is below the exercise price, $S_t < K$, and sets up a covered position as soon as the call is in–the–money, $S_t > K$.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_6,

Table 6.1 Parameters for the European call

Strike price K	50.00
Time to maturity τ	20 weeks = 0.3846
Riskfree rate r p.a.	0.05
Annualized stock volatility σ	0.20
Number of shares	100,000

Table 6.2 Stock price movement

t	0	1	2	3	4	5	6	7	8	9	10
S_t	49.00	49.75	52.00	50.00	48.37	48.25	48.75	49.62	48.25	48.25	51.12
$C(*10^5)$	2.40	2.72	4.07	2.69	1.75	1.61	1.75	2.10	1.34	1.25	2.67
t	11	12	13	14	15	16	17	18	19	20	
S_t	51.50	49.87	49.87	48.75	47.50	48.00	46.25	48.12	46.62	48.12	
$C(*10^5)$	2.82	1.69	1.56	0.91	0.40	0.41	0.06	0.18	0.00	0.00	

Exercise 6.1 (European Call Option). *Consider the European call detailed in Table 6.1 and the corresponding stock price movement given in Table 6.2. We assume, that there are no transaction costs nor opportunity costs arising from binding capital. However bear in mind that in this example it's only possible to sell or buy stocks at* $K \pm \delta$*.*

(a) Calculate the BS price of the call and the cost of hedging.
(b) Modify your quantlet for $S_0 = 51$*, i.e. the call is in the money.*

(a) The solution is provided in Table 6.3, where a dummy variable indicates the need of a hedge strategy and the resulting consequences, e.g. the number of shares purchased, the costs of buying shares, the revenues due to selling shares and the cumulative costs.

When the stock price crosses $K = 50$ at time point 2, the open position is hedged at a price of $S_t = 52$, resulting in costs of 5,200,000. In the next period, as $S_t \leq K$, we close the position at a price of $S_t = 50$ and receive 5,000,000. These transactions results in a loss of 2 per share, i.e. a loss of 200,000. We continue the stop- loss strategy, checking every period whether $S_t \geq K$ and $S_t \leq K$ respectively.

At the exercise point $T = 20$, if $S_t \geq K$ we will hold 100,000 shares bought at $S_t \geq K$ and will be able to serve the call option for which we receive K units per share. If $S_t \leq K$ the option will not be exercised and we hold no shares, as we will have sold them at some $S_t \leq K$.

SFSstoploss

(b) The solution is provided in Table 6.4. Note the difference in the final costs compared to Table 6.3

SFSstoploss

Table 6.3 Solution for Exercise 6.1 (a)

Time	Stock price	Hedge strategy	Shares purchased	Cost of shares	Revenue of shares	Cumulative costs
0	49.00	0	0	0	0	0
1	49.75	0	0	0	0	0
2	52.00	1	100,000	5,200,000	0	5,200,000
3	50.00	0	0	0	5,000,000	200,000
4	48.37	0	0	0	0	200,000
5	48.25	0	0	0	0	200,000
6	48.75	0	0	0	0	200,000
7	49.62	0	0	0	0	200,000
8	48.25	0	0	0	0	200,000
9	48.25	0	0	0	0	200,000
10	51.12	1	100,000	5,112,000	0	5,312,000
11	51.50	1	0	0	0	5,312,000
12	49.87	0	0	0	4,987,000	325,000
13	49.87	0	0	0	0	325,000
14	48.75	0	0	0	0	325,000
15	47.50	0	0	0	0	325,000
16	48.00	0	0	0	0	325,000
17	46.25	0	0	0	0	325,000
18	48.12	0	0	0	0	325,000
19	46.62	0	0	0	0	325,000
20	48.12	0	0	0	0	325,000

Exercise 6.2 (Stop-Loss Strategy). *Check the performance measure of a Stop-Loss strategy for an increasing hedging frequency Δt:*

(a) Consider the stock outlined in Table 6.1 and simulate $m = 500$ stock paths starting from $S_0 = 49.00$.
(b) Calculate, via a quantlet, the costs Λ_m for applying the Stop-Loss strategy over all m. Thereby set $\Delta t = 5$.
(c) Calculate the variance υ_Λ^2 of these costs.
(d) Calculate the following performance measure

$$L = \frac{\sqrt{\upsilon_\Lambda^2}}{C(S_0, T)}$$

(e) Modify your quantlet to calculate the performance measures L for $\Delta t = \{4, 2, 1, 0.5, 0.25\}$.

Compare these differing values of L, as it has exemplary been done in Table 6.5.

Table 6.4 Solution for Exercise 6.1 (b)

Time	Stock price	Hedge strategy	Shares purchased	Cost of shares	Revenue of shares	Cumulative costs
0	51.00	1	100,000	5,100,000	0	5,100,000
1	49.75	0	0	0	4,975,000	125,000
2	52.00	1	100,000	5,200,000	0	5,325,000
3	50.00	0	0	0	5,000,000	325,000
4	48.37	0	0	0	0	325,000
5	48.25	0	0	0	0	325,000
6	48.75	0	0	0	0	325,000
7	49.62	0	0	0	0	325,000
8	48.25	0	0	0	0	325,000
9	48.25	0	0	0	0	325,000
10	51.12	1	100,000	5,112,000	0	5,437,000
11	51.50	1	0	0	0	5,437,000
12	49.87	0	0	0	4,987,000	450,000
13	49.87	0	0	0	0	450,000
14	48.75	0	0	0	0	450,000
15	47.50	0	0	0	0	450,000
16	48.00	0	0	0	0	450,000
17	46.25	0	0	0	0	450,000
18	48.12	0	0	0	0	450,000
19	46.62	0	0	0	0	450,000
20	48.12	0	0	0	0	450,000

Table 6.5 Exemplary values for Exercise 6.2. `SFShullhedgeratio`

Δt	5	4	2	1	0.5	0.25
L	1.041	0.928	0.902	0.818	0.777	0.769

Exercise 6.3 (Delta Ratio). *Calculate the value of $\Delta = \frac{\partial C}{\partial S}$, the ratio of change of the option price with respect to the underlying stock price, for a European call with strike price K and maturity T.*

The BS formula for a European call may be written as follows:

$$C(S,\tau) = S\Phi(d_1) - K\exp(-r\tau)\Phi(d_2),$$

where we use the abbreviations

$$\begin{aligned} d_1 &= y + \sigma\sqrt{\tau}, \\ d_2 &= y \\ y &= \frac{\log S/K + \{r - \sigma^2/2\}\,\tau}{\sigma\sqrt{\tau}}. \end{aligned}$$

A simple derivative of the formula yields

$$\frac{\partial C(S,\tau)}{\partial S} = \Phi(d_1) + S\frac{\partial \Phi(d_1)}{\partial S} - K\exp(-r\tau)\frac{\partial \Phi(d_2)}{\partial S},$$

where we again assume $b = r$ to clarify the calculations.

By inserting the following intermediate results

$$\begin{aligned}
\frac{\partial \Phi(d_1)}{\partial S} &= \varphi(d_1)\frac{\partial d_1}{\partial S} \\
\frac{\partial \Phi(d_2)}{\partial S} &= \varphi(d_2)\frac{\partial d_2}{\partial S} \\
\frac{\partial d_1}{\partial S} = \frac{\partial d_2}{\partial S} &= \left(S\sigma\sqrt{\tau}\right)^{-1} \\
\varphi(d_2) &= \varphi(d_1)\exp(r\tau)\frac{S}{K}
\end{aligned}$$

we obtain the following delta ratio:

$$\begin{aligned}
\frac{\partial C(S,\tau)}{\partial S} &= \Phi(d_1) + S\varphi(d_1)\left(S\sigma\sqrt{\tau}\right)^{-1} \\
&\quad - K\exp(-r\tau)\varphi(d_1)\exp(r\tau)\frac{S}{K}\left(S\sigma\sqrt{\tau}\right)^{-1} \\
&= \Phi(d_1).
\end{aligned}$$

Exercise 6.4 (Delta Hedging). *Calculate the delta ratio $\Delta = \frac{\partial C(S,\tau)}{\partial S}$ for the European call stated in Exercise 6.1. Furthermore think of the consequences resulting from the delta hedge, e.g. the amount of shares to be held and the cumulative costs for holding these stocks across t.*

To calculate the delta ratio, which changes across time t due to the stock price S_t and the declining time to maturity τ, we use the result from Exercise 6.3, namely

$$\frac{\partial C(S,\tau)}{\partial S} = \Phi\left\{\frac{\log S/K + \left(r - \sigma^2/2\right)\tau}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right\}.$$

With this ratio, we can calculate the number of shares we need to hold at any t and the resulting costs for holding these shares across time. The solution is provided in Table 6.6.

Exercise 6.5 (Delta Neutral Position). *A bank sold 3,000 Calls and 2,500 Puts with the same maturity and exercise price on the same underlying stock. Δ Call is 0.42 and Δ Put is −0.38. How many stocks should the company sell or buy in order to have a delta neutral position?*

Table 6.6 Solution to Exercise 6.4. `SFSdeltahedging`

Time	Stock price	Delta Hedging	Shares purchased	Cost of shares	Revenue of shares	Cumulative costs
0	49.000	0.5216	52,160	2,555,840.0	0.0	2,555,840.0
1	49.750	0.5675	56,750	228,352.5	0.0	2,784,192.5
2	52.000	0.7051	70,510	715,520.0	0.0	3,499,712.5
3	50.100	0.5861	58,610	0.0	596,190.0	2,903,522.5
4	48.375	0.4587	45,870	0.0	616,297.5	2,287,225.0
5	48.250	0.4429	44,290	0.0	76,235.0	2,210,990.0
6	48.750	0.4751	47,510	156,975.0	0.0	2,367,965.0
7	49.625	0.5397	53,970	320,577.5	0.0	2,688,542.5
8	48.250	0.4197	41,970	0.0	579,000.0	2,109,542.5
9	48.250	0.4105	41,050	0.0	44,390.0	2,065,152.5
10	51.125	0.6581	65,810	1,265,855.0	0.0	3,331,007.5
11	51.500	0.6918	69,180	173,555.0	0.0	3,504,562.5
12	49.875	0.5420	54,200	0.0	747,127.5	2,757,435.0
13	49.875	0.5376	53,760	0.0	21,945.0	2,735,490.0
14	48.750	0.3998	39,980	0.0	671,775.0	2,063,715.0
15	47.500	0.2362	23,620	0.0	777,100.0	1,286,615.0
16	48.000	0.2615	26,150	121,440.0	0.0	1,408,055.0
17	46.250	0.0619	6,190	0.0	923,150.0	484,905.0
18	48.120	0.1818	18,180	576,958.8	0.0	1,061,863.8
19	46.620	0.0067	670	0.0	816,316.2	245,547.6
20	48.120	0.0000	0	0.0	32,240.4	213,307.2

Let's call the number of stocks needed for a delta neutral position x. The Δ of the stock is 1. From Franke et al. (2011, p. 114) we will have

$$\begin{aligned} 0.42 \cdot (-3000) - 0.38 \cdot (-2500) + x \cdot 1 &= 0 \\ x &= 310 \end{aligned}$$

So, the bank will have a delta neutral position when it buys 310 stocks.

Exercise 6.6 (Gamma and Delta Hedging). *Mrs. Ying Chen already has a portfolio with $\Delta = -300$ and $\Gamma = 250$. She wants to make her portfolio Δ and Γ neutral. In the market, she can buy/sell stocks and call options to achieve this. She calculated the Δ and Γ of the call option as 0.55 and 1.25 respectively. What should Mrs. Chen do in order to realize her goal?*

Let us call the additional number of calls and stocks that Mrs. Chen will need as x and y respectively. For a Δ and Γ hedged portfolio, we can write the following equations from the corresponding Gamma and Delta Hedging equations:

$$1.25 \cdot x + 250 = 0$$

$$0.55 \cdot x + y \cdot 1 - 300 = 0$$

After solving this set of equations, we see, that Mrs. Chen should buy $y = 410$ stocks and short sell $x = 200$ call options in order to achieve a Δ and Γ neutral position.

Exercise 6.7 (Option Pricing with Black-Scholes Model). *Consider a European call option on a stock when there are ex-dividend dates in 3 and 6 months. The dividend on each ex-dividend date is expected to be 1, the current price is 80, the exercise price is 80, the volatility is 25 % per annum, the annually risk free rate is 7 %, the time to maturity is 1 year, calculate the option price using the Black-Scholes model.*

The present value of the expected dividends can be calculated as

$$\exp(-0.25 \cdot 0.07) + \exp(-0.5 \cdot 0.07) = 1.9483.$$

We deduct this present value of the dividends from the stock price at $t = 0$ to arrive at the purely random component of the stock price:

$$\begin{aligned} S_0 &= 80 - 1.9483 \\ &= 78.0517. \end{aligned}$$

Afterwards, as a first step, we calculate $\Phi(y)$ and $\Phi(y + \sigma\sqrt{\tau})$ according to the formulas given in Exercise 6.3:

$$\begin{aligned} \Phi(y) &= \Phi\left[\frac{\log(\frac{78.0517}{80}) + \left\{0.07 - (0.25)^2/2\right\} \cdot 1}{0.25 \cdot \sqrt{1}}\right] \\ &= \Phi(0.0564) \\ &= 0.5225 \\ \Phi(y + \sigma\sqrt{\tau}) &= \Phi\left(0.0564 + 0.25 \cdot \sqrt{1}\right) \\ &= \Phi(0.3064) \\ &= 0.6203 \end{aligned}$$

We now possess all information we need to calculate the price of the call:

$$\begin{aligned} C(78.0517, 1) &= 78.0517 \cdot 0.6203 - 80 \cdot \exp(-0.07 \cdot 1) \cdot 0.5225 \\ &= 9.4462 \end{aligned}$$

SFSBSCopt1

Exercise 6.8 (Delta, Gamma and Theta of Portfolio). *An investor longs a 3 months maturity call option with* $K = 220$*, and shorts a half year call with* $K = 220$*. Currently,* $S_0 = 220$*, risk free rate* $r = 0.06$ *(continuously compounded), and* $\sigma = 0.25$*. The stock pays no dividends.*

(a) Calculate the Δ of the portfolio.
(b) Calculate the Γ of the portfolio.
(c) Calculate the Θ of the portfolio.

(a) We calculate $\Delta = \Phi(y + \sigma\sqrt{\tau})$ for both calls separately and sum up the results to obtain the overall delta of the portfolio:

Call to be longed:
$$\Delta_1 = 0.6018$$
Call to be shorted:
$$\Delta_2 = -0.5724$$
The delta of the portfolio is therefore
$$\begin{aligned} \Delta_1 + \Delta_2 &= 0.6018 + (-0.5724) \\ &= 0.0294 \end{aligned}$$

(b) To obtain the portfolio's gamma we calculate
$$\Gamma = \frac{1}{\sigma S\sqrt{\tau}}\varphi\left(y + \sigma\sqrt{\tau}\right)$$
for each call:

Call to be longed:
$$\Gamma_1 = 0.0099$$
Call to be shorted:
$$\Gamma_2 = -0.0142$$
This leads to the following overall gamma:
$$\begin{aligned} \Gamma_1 + \Gamma_2 &= 0.0099 + (-0.0142) \\ &= -0.0043 \end{aligned}$$

(c) By the derivative of the BS formula with respect to t
$$\begin{aligned} \Theta &= \frac{\partial C(S,\tau)}{\partial t} \\ &= -\frac{\sigma S}{2\sqrt{\tau}}\varphi\left(y + \sigma\sqrt{\tau}\right) - rK\exp(-r\tau)\Phi(y) \end{aligned}$$

we can calculate the theta for each call:

Call to be longed:

$$\Theta_1 = -21.827$$

Call to be shorted:

$$\Theta_2 = -28.279$$

The theta of the portfolio may then be calculated as:

$$\begin{aligned}\Theta_1 - \Theta_2 &= -21.827 + (-28.279)\\ &= -50.106\end{aligned}$$

SFSgreeks

Exercise 6.9 (Collar Portfolio). *Mr. Wang constructed a collar portfolio, which was established by buying a share of a stock for 15 JPY, buying a 1-year put option with exercise price 12.5 JPY, and short selling a 1-year call option with exercise price 17.5 JPY. If the BS model holds, based on the volatility of the stock, Mr. Wang calculated that for a strike price of 12.5 JPY and maturity of 1 year,* $\Phi(y+\sigma\sqrt{\tau}) = 0.63$, *whereas for the exercise price of 17.5 JPY,* $\Phi(y+\sigma\sqrt{\tau}) = 0.32$ *where we use* y *as an abbreviation for*

$$y = \frac{\log S/K + \left(b - \sigma^2/2\right)\tau}{\sigma\sqrt{\tau}}. \tag{6.4}$$

(a) Draw a payoff graphic of this collar at the option expiration date.
(b) If the stock price increases by 1 JPY, what will Mr. Wang gain or lose from this portfolio?
(c) What happens to the delta of the portfolio if the stock price becomes very high or very low?

(a) The payoff can be drawn as in the Fig. 6.1.
(b) The resulting consequences for Mr. Wang's portfolio may be calculated via the delta of the portfolio, which indicates any losses or gains. In detail, the value of the stock will raise by 1 JPY, the loss on the long put and on the short call will amount to 0.37 and 0.32 JPY respectively. For the whole portfolio, this leads to a gain of 0.31 JPY, as stated in Table 6.7.
(c) For very large stock prices: the delta of the collar approaches zero, because both $\pm\Phi(y+\sigma\sqrt{\tau})$ approach 1. The value of the portfolio is simply the present value of the exercise price of the call, and is unaffected by small changes in the stock price.

For very small stock prices: as stock price approaches zero, the delta also approaches zero, because both $\pm\Phi(y+\sigma\sqrt{\tau})$ terms approach 0. The value of the portfolio is simply the exercise price of the put, and is unaffected by small changes in the stock price.

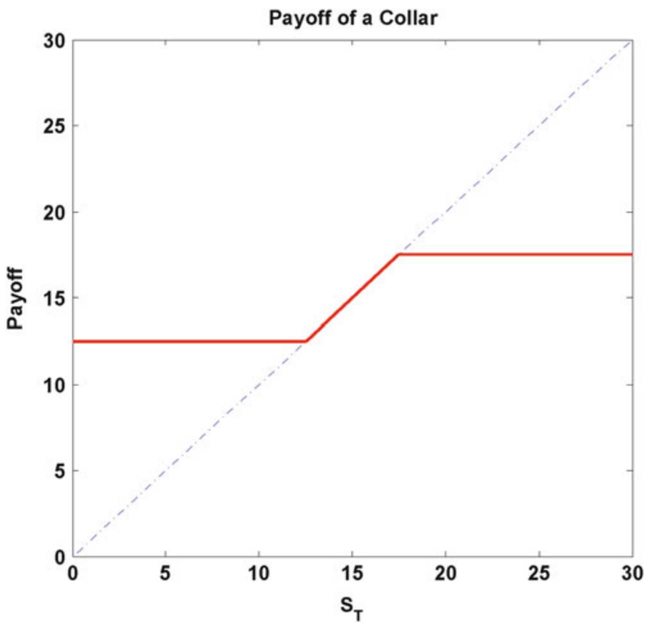


Fig. 6.1 Payoff of a collar. SFSpayoffcollar

Table 6.7 Delta of the collar

Action	Delta
Long a stock	1.00
Long a put	$\Phi(y+\sigma\sqrt{\tau})-1=-0.37$
Short a call	$-\Phi(y+\sigma\sqrt{\tau})=-0.32$
Portfolio	0.31

Table 6.8 Portfolio on a particular stock

Type	Position	Delta	Gamma
Call	−2,000	0.5	2.5
Call	+1,500	0.7	0.7
Put	−4,000	−0.4	1.1

Exercise 6.10 (Gamma and Delta Neutral Positions).

A financial institution has the portfolio given in Table 6.8 of OTC options on a particular stock.

A traded call option is available which has a delta $\Delta = 0.3$ *and a gamma* $\Gamma = 1.8$. *What position in that traded option and underlying stock would make the portfolio both gamma neutral and delta neutral?*

We can calculate the current position of portfolio:

$$\begin{aligned}\varDelta &= -2000 \cdot 0.5 + 1500 \cdot 0.7 - 4000 \cdot (-0.4)\\ &= -1000 + 1050 + 1600\\ &= 1650\\ \varGamma &= -2000 \cdot 2.5 + 1500 \cdot 0.7 - 4000 \cdot 1.1\\ &= -5000 + 1{,}050 - 4{,}400\\ &= -8350\end{aligned}$$

In order to become Gamma-neutral according to Franke et al. (2011, Sect. 6.4), one should buy traded options:

$$\frac{8350}{1.8} = 4638.89.$$

This will make the delta of the position equal to

$$1650 + 4638.89 \cdot 0.3 = 3041.67.$$

Therefore, to become delta neutral we need to sell 3,041.67 shares.

Exercise 6.11 (Hypothetical Call Option). *Knowing that the current price of oil is 100 EUR per barrel, a petrochemical firm PetroCC plans to buy a call option on oil with strike price = 100 EUR. The volatility of oil prices is 10 % per month, and the risk-free rate is 3 % per month.*

(a) What is the value of the 4-month call option?
(b) Suppose that instead of buying 4-month call options on 100,000 barrels of oil, the firm will synthetically replicate a call position by buying oil directly now, and delta-matching the hypothetical call option. How many barrels of oil should it buy?
(c) If oil prices increase by 1 % after the first day of trading, how many barrels of oil should it buy or sell?

(a) We can calculate the call option price using BS formula directly:

$$C(S,\tau) = S\varPhi(y + \sigma\sqrt{\tau}) - \exp(-r\tau)K\varPhi(y)$$

where $S = 100$, $K = 100$, $\sigma = 0.1$, $b = r = 0.03$ and $\tau = 4$.
Then,

$$\begin{aligned}y + \sigma\sqrt{\tau} &= \frac{\log S/K + (b + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}\\ &= \frac{0 + (0.03 + 0.10^2/2) \cdot 4}{0.10 \cdot \sqrt{4}}\\ &= 0.7000\end{aligned}$$

$$\begin{aligned}
y &= 0.7000 - \sigma\sqrt{\tau} \\
&= 0.5000 \\
\Phi(y + \sigma\sqrt{\tau}) &= 0.7580 \\
\Phi(y) &= 0.6915 \\
C(S,\tau) &= S\Phi(y + \sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(y) \\
&= 100 \cdot 0.7580 - \exp(-0.03 \cdot 4) \cdot 100 \cdot 0.6915 \\
&= 14.4695
\end{aligned}$$

So the price of 4-month call option should be 14.4695 EUR.

(b) Since the delta of the call option is,

$$\Delta = \frac{\partial C}{\partial S} = \Phi(y + \sigma\sqrt{\tau}) = 0.7580$$

and the delta of the stock is 1.
So, we should buy 75,800 barrels of oil, which can provide the same delta value as the call options.

(c) If the oil price increases by 1 %, then the $y + \sigma\sqrt{\tau}$ also increases, so as the delta.

$$\begin{aligned}
y + \sigma\sqrt{\tau} &= \frac{\log S/K + (b + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} \\
&= \frac{0.01 + (0.03 + 0.10^2/2) \cdot 4}{0.10 \cdot \sqrt{4}} \\
&= 0.8000 \\
\Phi(0.800) &= 0.7881
\end{aligned}$$

The delta increases $0.7881 - 0.7580 = 0.0301$, so one should buy an additional 3,010 barrels of oil.

Exercise 6.12 (Implied Volatility and Delta Neutrality). *There are two calls on the same stock with the same time to maturity (1 year) but different strike price. Option A has a strike price $K_1 = 10$ USD, while option B has a strike price $K_2 = 9.5$ USD. The current stock price is $S_t = 10$ USD. The stock does not pay dividend. Risk free rate is 3 %. One applies the Black-Scholes equation as the option pricing model. Yet despite that fact that one is confident that the appropriate volatility of the stock is 0.15 p.a. One observes option A selling for 0.8 USD and option B selling for 1 USD.*

(a) Is the implied volatility of Option A more or less than 15 %? What about that of option B?

(b) Determine a delta-neutral position in the two calls that will exploit their apparent mispricing. Use $\sigma = 0.15$ *to compute Delta.*

(a) We can calculate the call option price using BS formula directly:
For option A:

$$\begin{aligned}
y + \sigma\sqrt{\tau} &= \frac{\log S/K + (b + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} \\
&= \frac{0 + (0.03 + 0.15^2/2)\cdot 1}{0.15\cdot\sqrt{1}} \\
&= 0.28 \\
y &= 0.28 - \sigma\sqrt{\tau} \\
&= 0.13 \\
\Phi(y + \sigma\sqrt{\tau}) &= 0.6103 \\
\Phi(y) &= 0.5517 \\
C_A(S,\tau) &= S\Phi(y + \sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(y) \\
&= 10\cdot 0.6103 - \exp(-0.03\cdot 1)\cdot 10\cdot 0.5517 \\
&= 0.75 < 0.8
\end{aligned}$$

For option B:

$$\begin{aligned}
y + \sigma\sqrt{\tau} &= \frac{\log S/K + (b + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} \\
&= \frac{0.051 + (0.03 + 0.15^2/2)\cdot 1}{0.15\cdot\sqrt{1}} \\
&= 0.62 \\
y &= 0.62 - \sigma\sqrt{\tau} \\
&= 0.47 \\
\Phi(y + \sigma\sqrt{\tau}) &= 0.7324 \\
\Phi(y) &= 0.6808 \\
C_B(S,\tau) &= S\Phi(y + \sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(y) \\
&= 10\cdot 0.7324 - \exp(-0.03\cdot 1)\cdot 9.5\cdot 0.6808 \\
&= 1.05 > 1
\end{aligned}$$

For Option A, the real price is higher than the Black-Scholes value when using 0.15 as the volatility. For option B, the price is lower. Thus, for option A, the implied volatility should be higher than 0.15, and the implied volatility of option B is lower than 0.15.

(b) From task (a), it is clear that option A is overvalued, and option B is undervalued. Thus, we can long option B and short option A to do arbitrage. For every option A we short, we should long $\Delta_A/\Delta_B = 0.6103/0.7324 = 0.8333$ option B.

Exercise 6.13 (Implied Volatility). *E-Tech Corp. stock sells for 80 EUR and pays no dividends. A 6-month call option with exercise price 80 EUR is priced at 7.23 EUR, while a 6-month call option with exercise price 90 EUR is priced at 5.38. The risk-free interest rate is 8 % per year.*

(a) What is the implied volatility of these two options?
(b) Based on the above information, construct a profitable trading strategy.

(a) From Black-Scholes formula, we have:

$$C(S,\tau) = S\Phi(y+\sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(y) \tag{6.5}$$

$$y+\sigma\sqrt{\tau} = \frac{\log S/K + (b+\sigma^2/2)\tau}{\sigma\sqrt{\tau}} \tag{6.6}$$

After substituting data in this example, solve the equations w.r.t. σ, we have the implied volatility of the call option with exercise price 80 EUR is 25 %; the implied volatility of the call option with exercise price 90 EUR is 35 %. (The calculations in this question can also been done in the DerivaGem Options Calculator software of John Hull.)

(b) Based on implied volatilities, the call option with exercise price 90 EUR is overpriced. Thus we can exploit this by constructing a delta-neutral position: buy the call with exercise price of 80 EUR and write the call with exercise price of 90 EUR.

To calculate delta, we need to know σ. One strategy is to use the middle point of 25 and 35 %, i.e., $\sigma = 35$ %. When $\sigma = 35$ %, the delta of the call with exercise price of 80 EUR is 0.616, and the delta of the call with exercise price of 90 EUR is 0.397. Thus the delta-neutral proportion is 0.616/0.397 = 1.552. So we buy one call with exercise price of 80 EUR, and write 1.552 calls with exercise price of 90 EUR. The net position is delta neutral, but since $7.32 - 1.552 \times 5.38 = -1.03$, there is an initial cash inflow (profit) of 1.03 EUR.

Exercise 6.14 (Greeks). *The Black-Scholes price of a call option with strike price K, maturity T is defined as follows in $t \in [0,T)$ at a stock price x:*

$$v(x,T-t) = x\Phi\{d_+(x,T-t)\} - K\exp\{-r(T-t)\}\Phi\{d_-(x,T-t)\}$$

$r \in \mathbb{R}$ denotes the riskless interest rate and $\sigma > 0$ the volatility of the stock. The function d_+ and d_- are given by

$$d_{\pm}(x, r) = \{\log x/K + (r \pm \sigma^2/2)\tau\}/\sigma\sqrt{\tau}$$

Calculate the "Greeks"

(a) $\Delta = \frac{\partial}{\partial x} v(x, T-t)$
(b) $\Gamma = \frac{\partial^2}{\partial x^2} v(x, T-t)$
(c) $\Theta = \frac{\partial}{\partial t} v(x, T-t)$

and verify that v solves the partial differential equation

$$\left(\frac{\sigma^2}{2}x^2\frac{\partial^2}{\partial x^2} + rx\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right)v(x,t) = rv(x,t) \text{ on } (0,\infty)\times(0,\infty)$$

and satisfies the boundary condition

$$v(x, T-t) \to (x-K)^+ \text{ for } t \to T$$

From the Black-Scholes formula,

$$v(x, T-t) = x\Phi\{d_+(x, T-t)\} - K\exp\{-r(T-t)\}\Phi\{d_-(x, T-t)\}$$

where Φ is the distribution function of the standard normal distribution and $(T-t) = \tau$ the time to maturity.
Recall (Exercise 6.3), the ratio of change of the option price with respect to the underlying stock price (Delta, $\Delta = \frac{\partial v}{\partial x}$) can be expressed as:

(a)

$$\begin{aligned}\frac{\partial}{\partial x}v(x, T-t) = \Phi\{d_+(x, T-t)\} + x\varphi\{d_+(x, T-t)\}\frac{\partial}{\partial x}d_+(x, T-t) - \\ \varphi\{d_-(x, T-t)\}\frac{\partial}{\partial x}d_-(x, T-t)\end{aligned} \tag{6.7}$$

Note that,

$$\begin{aligned} d_+(x,\tau) - d_-(x,\tau) &= \left\{\left(r + \frac{\sigma^2}{2}\right)\tau - \left(r - \frac{\sigma^2}{2}\right)\tau\right\}/\sigma\sqrt{\tau} = \sigma\sqrt{\tau} \\ \frac{\partial}{\partial x}d_\pm(x,\tau) &= k/(\sigma\sqrt{\tau}xk) = 1/x\sigma\sqrt{\tau} \\ \varphi(d_-(x,\tau)) &= \exp\{-d^2/2 - (x,\tau)\}/2\pi \end{aligned}$$

$$= \exp\{-(d_+(x,\tau) - \sigma\sqrt{\tau})^2/2\}/\sqrt{2\pi}$$
$$= \exp\{-(d^2 + (x,\tau) - 2d + (x,\tau)\sigma\sqrt{\tau} + \sigma^2\tau)/2\}/\sqrt{2\pi}$$
$$= \varphi\{d_+(x,\tau)\}\exp\{\sigma\sqrt{\tau}d_+(x,\tau) - \sigma^2\tau/2\}$$
$$= \varphi\{d_+(x,\tau)\}\exp(\log x/K + r\tau)$$
$$= \varphi\{d_+(x,\tau)\}x/K\exp(r\tau)$$

By substitution in Eq. 6.7,

$$\text{Delta: } \Delta = \frac{\partial \nu}{\partial x} = \Phi\{d_+(x,T-t)\} + x\varphi\{d_+(x,T-t)\}\frac{\partial}{\partial x}d_+(x,T-t) - \varphi\{d_-(x,T-t)\}\frac{\partial}{\partial x}d_-(x,T-t) = \Phi\{d_+(x,T-t)\}$$

(b) The ratio of change of the option Δ with respect to the underlying stock price (Gamma, $\Gamma = \frac{\partial \nu}{\partial \Delta}$) can be expressed as:

$$\Gamma = \frac{\partial \nu}{\partial \Delta} = \frac{\partial}{\partial x^2}\nu(x,T-t) = \frac{\partial}{\partial x}\Phi\{d_+(x,T-t)\}$$
$$= \varphi\{d_+(x,T-t)\}\frac{\partial}{\partial x}d_+(x,T-t)$$
$$= \frac{\varphi\{d_+(x,T-t)\}}{x\sigma\sqrt{T-t}}$$

(c) The ratio of change of the price of the underlying with respect to time (Theta, $\Theta = \frac{\partial \nu}{\partial t}$) can be expressed as:

$$\Theta = \frac{\partial}{\partial t}\nu(x,T-t) = x\varphi\{d_+(x,T-t)\}\frac{\partial}{\partial t}d_+(x,T-t)$$
$$-kr\exp\{-r(T-t)\}\Phi\{d_-(x,T-t)\}$$
$$+\exp\{-r(T-t)\}\varphi\{d_-(x,T-t)\}\frac{\partial}{\partial t}d_-(x,T-t)$$
$$= x/k\varphi\{d_+(x,T-t)\}$$
$$= x\varphi\{d_+(x,T-t)\}\frac{\partial}{\partial t}(d_+d_-)(x,T-t) -$$
$$-kr\exp\{-r(T-t)\}\Phi\{d_-(x,T-t)\}$$
$$= -\frac{x\sigma}{d\sqrt{(T-t)}}\varphi\{d_+(x,T-t)\}$$
$$-kr\exp\{-r(T-t)\}\Phi\{d_-(x,T-t)\}$$

To verify that ν solves the partial differential equation, we need to show that the Black-Scholes option price model gives the same price as a model free no-arbitrage approach. Applying Itô's lemma:

$$d\nu(x,t) = \sigma x \frac{\partial \nu}{\partial x} dW_t + \left(\mu x \frac{\partial \nu}{\partial x} + \frac{\sigma^2}{2} x^2 \frac{\partial^2 \nu}{\partial x^2} + \frac{\partial \nu}{\partial t} \right) dt.$$

Consider a portfolio Π containing an option and $-\Delta$ units of the underlying stocks:

$$\Pi = \nu(x,t) - \Delta x$$

$$d\Pi = d\nu(x,t) - \Delta dx$$

$$d\Pi = d\nu(x,t) - \Delta(\mu x dt + \sigma x dW_t)$$

For $\Delta = \frac{\partial \nu(x,t)}{\partial x}$

$$d\Pi = \left(\frac{\sigma^2}{2} x^2 \frac{\partial^2 \nu}{\partial x^2} + \frac{\partial \nu}{\partial t} \right) dt$$

Now if Π was invested in riskless assets it would see a growth of $r\Pi dt$ in the interval of length dt. Then for a fair price we should have $d\Pi = r\Pi dt$.

$$r\Pi dt = \left\{ \frac{\sigma^2}{2} x^2 \frac{\partial^2 \nu(x,t)}{\partial x^2} + \frac{\partial \nu(x,t)}{\partial t} \right\} dt$$

Hence,

$$r\left\{ \nu - \frac{\partial \nu(x,t)}{\partial x} x \right\} = \frac{\sigma^2}{2} x^2 \frac{\partial^2 \nu(x,t)}{\partial x^2} + \frac{\partial \nu(x,t)}{\partial t}.$$

Re-arranging gives

$$\frac{\sigma^2}{2} x^2 \frac{\partial^2 \nu(x,t)}{\partial x^2} + rx \frac{\partial \nu(x,t)}{\partial x} + \frac{\partial \nu(x,t)}{\partial t} - r\nu(x,t) = 0,$$

satisfying $\nu(x,\tau) \to (x-K)^+$ for $t \to T$.

Exercise 6.15 (Black-Scholes Price of a Call Option and Vega).

(a) Show that for $x, y > 0$ with $x \neq t$, the following holds:

$$|\nu(x,t) - \nu(y,t)| < |x - y|,$$

$$\frac{\nu(x,t) - \nu(y,t)}{\nu(y,t)} > \frac{x-y}{y} \text{ for } x > y,$$

and

$$\frac{\nu(x,t) - \nu(y,t)}{\nu(y,t)} < \frac{x-y}{y} \text{ for } x < y$$

for the special cases, where $\nu(x,t)$ is the Black-Scholes price of a call option with stock price x, time to expiration of the option as t.

(b) Show that the "Vega" $= \frac{\partial}{\partial\sigma}\nu(x,t)$ is always positive and calculate the value x where "Vega" is maximal.

(a) For $\nu(x, T-t)$ (see, question Exercise 6.14), let time to maturity $\tau = T - t$. It holds that

$$\frac{\partial}{\partial x^2}\nu(x,\tau) = \varphi\{d_+(x,\tau)\}/x\sigma\sqrt{\tau} > 0,$$

hence

$$\nu(x,\tau) \text{ is strictly convex on } (0,\infty)$$

Following Lipschitz condition, $|\nu(x,\tau)-\nu(y,\tau)| < |x-y|$ for $x, y > 0, x \neq y$ $|\nu(x,\tau) - \nu(y,\tau)| = |\nu_x(z,\tau)||x-y|$

It holds $\forall\ z \in (0, y)$:

$$\frac{\nu(x,\tau)-\nu(y,\tau)}{x-y} > \frac{\nu(y,\tau)-\nu(z,\tau)}{y-z},$$

therefore

$$\frac{\nu(x,\tau)-\nu(x,\tau)}{x-y} > \frac{\nu(y,\tau)}{y}$$

Also for $x < y$,

$$\frac{\nu(x,\tau)-\nu(y,\tau)}{\nu(y,\tau)} < \frac{x-y}{y},$$

if and only if

$$\frac{y-x}{x} < \frac{\nu(y,\tau)-\nu(x,\tau)}{\nu(x,\tau)}$$

(b) The "Vega" is the rate of change the option with respect to the volatility. For a European call option on a non-dividend-paying stock,

$$\text{Vega: } \frac{\partial\nu(x,\tau)}{\partial\sigma} = \varphi\{d_+(x,\tau)\}\sqrt{\tau}$$
$$= \frac{x\sqrt{\tau}}{\sqrt{2\pi}}\exp(-d_1^2/2),$$

where

$$d_1 = \frac{\log x/K + \left(r - \sigma^2/2\right)\tau}{\sigma\sqrt{\tau}},$$

which is always positive. Let $V =$ Vega, then

$$V' = \frac{\sqrt{\tau}}{\sqrt{2\pi}} \exp(-d_1^2/2) \left(1 - \frac{d_1}{\sigma\sqrt{\tau}}\right).$$

Solving for x in V'', the value for which the Vega reaches maximum is obtained:

$$V'' = \frac{1}{\sqrt{2\pi}} \exp(-d_1^2/2) \left\{-d_1 + \frac{(d_1^2 - 1)}{\sigma\sqrt{\tau}}\right\},$$

with maximum Vega at $x = k \exp\{(\sigma^2 - r)\tau/2\}$

Exercise 6.16 (Price of Risk, Stochastic Process and Girsanov Transformation). *In the Black-Scholes model, the stock price is modelled by*

$$dS(t) = S(t)\{\mu dt + \sigma dW(t)\},$$

where μ denotes the drift, σ the volatility and Wt, the standard Brownian motion.

(a) The market price of risk is defined as excess return. How is this defined in the Black-Scholes framework?
(b) Give the explicit form of the stochastic process $S(t)$
(c) Suppose now that a class of parametrized class of equivalent probabilities Q are introduced via the Girsanov transformation:

$$W^\theta(t) = W(t) - \int_0^t \theta(u)du$$

where θ is a real valued, bounded continuous function. By using the Girsanov Theorem there exists an equivalent probability measure denoted Q^θ so that $W^\theta(t)$ is a Brownian motion for t. Show the dynamics of $S(t)$ under Q^θ

(a) The market price of risk is the rate of extra return above r per unit risk. In the Black-Scholes model, the stock price is modelled by

$$dS(t) = S(t)\{\mu dt + \sigma dW(t)\}.$$

Under an equivalent risk neutral measure the model can be expressed as

$$\frac{dS(t)}{S(t)} = (\mu - r)dt + \sigma dW(t).$$

By setting $X_t = \frac{\mu - r}{\sigma}$, the market price of risk:

$$\frac{dS(t)}{S(t)} = \sigma\{X_t dt + dW(t)\},$$

where r is the constant risk-free interest rate.

(b) By Itô's formula,

$$d\log S(t) = \left(\mu - \sigma^2/2\right)dt + \sigma dW(t),$$

so that $S(t)$ satisfies the Black-Scholes model, if and only if

$$S(t) = S(0)\exp\left\{\left(\mu - \sigma^2/2\right)t + \sigma W(t)\right\}$$

(c) We construct under Q^θ a martingale price process:

$$dS(t) = S(t)\{(\mu - r)dt + \sigma dW(t)\},$$

where (by Girsanov transformation)

$$W^\theta(t) = W(t) - \int_0^t \theta(u)du$$

is a Q^θ−Brownian motion. Applying the Radon- Nicodym derivative, $\zeta_t = \frac{dQ}{dQ^\theta}$, gives:

$$\zeta_t = \exp\left\{-\int_0^t \theta(u)dW(u) - \frac{1}{2}\int_0^t \theta^2(u)du\right\}.$$

From $dS(t) = S(t)\{(\mu - r)dt + \sigma dW(t)\}$ and by definition of $W^\theta(t)$, it holds that

$$dS(t) = S(t)\sigma dW^\theta(t).$$

Applying Itô's lemma the dynamics of the stochastic process is expressed as

$$S(t) = S(0)\exp\left\{-\int_0^t \sigma dW^\theta(u) - \frac{1}{2}\int_0^t \sigma^2 d(u)\right\}.$$

Chapter 7
Binomial Model for European Options

Binomialni model za europske opcije
Najveći je rizik ne riskirati!
The greatest risk is not to take risk!

For a large range of options such as the American options the boundary conditions of the Black-Scholes differential equation are too complex to solve analytically. Therefore, one relies on numerical price computation. The best known method is to approximate the stock price process by a discrete time stochastic process, or, as in the approach followed by Cox, Ross, Rubinstein, to model the stock price process as a discrete time process from the start. The binomial model is a convenient tool for pricing European options.

Exercise 7.1 (Value of a Call Option). *Assume a call option with exercise price* $K = 8$ *at* $T = 2$*. We are now at* $t = 0$*. The current price of the stock is* 10*. For the first period, the stock market is expected to be very variable, and the underlying stock's price is expected to increase or decrease by* 20 %*. For the second year, a more stable market is expected and the stock price is expected to increase or decrease by* 10 %*. Assume that the risk-free rate is zero. What is the value of this call option? The intrinsic value at* T *is denoted as* C_T*.*

$t = 0$	$t = 1$	$t = 2$	C_T
		13.2	5.2
	12		
		10.8	2.8
10			
		8.8	0.8
	8		
		7.2	0.0

DOI 10.1007/978-3-642-33929-5_7,

The price movement of the stock and the option price at maturity can be seen in the table above. To calculate the current value of the option, we should first find option prices at the end of period 1 by making use of the intrinsic values of the call option at the end of period 2. Afterwards, we will be able to price the call option for period 0 following a similar procedure.

We use the martingale measure approach to price the call option. Let us consider the upper part of the stock tree first:

$t = 1$	$t = 2$	C_T
	13.2	5.2
12		
	10.8	2.8

The probability associated with the movements from $t = 1$ to $t = 2$ is:

$q \times 13.2 + (1 - q) \times 10.8 = 12$, therefore $q = 0.5$. Hence, the corresponding option price at the end of the period 1 is:

$$C_1 = 0.5 \times 5.2 + 0.5 \times 2.8 = 4, \text{ as} \exp(r \times \triangle t) = 1 \text{ for } r = 0.$$

We repeat the whole procedure for the lower part of the tree:

$t = 1$	$t = 2$	C_T
	8.8	0.8
8		
	7.2	0.0

For the transition probability, it holds: $q \times 8.8 + (1 - q) \times 7.2 = 8$, therefore $q = 0.5$. The option price at the end of the period 1 in the lower part of the tree is: $C_1 = 0.5 \times 0.8 + 0.5 \times 0 = 0.4$.

Now, we can construct a stock tree with intrinsic values of the call option in the 1st period:

$t = 0$	$t = 1$	C_1
	12	4.0
10		
	8	0.4

We compute the transition probability from the stock prices:

$q \times 12 + (1 - q) \times 8 = 10$ gives $q = 0.5$, and afterwards the call option price at the $t = 0$: $C_0 = 0.5 \times 4 + 0.5 \times 0.4 = 2.2$.

Exercise 7.2 (Trinomial Process). *Assume that a stock's daily returns exhibit a trinomial process. With equal probabilities ($p = 1/3$), the stock's price either increases 3 or 2 %, or it decreases 4 % each day. What can be said about the price of this stock at the end of the year, assuming a $T = 260$ work days and an initial stock price of $X_0 = 100$?*

The stock price X_t follows a geometric trinomial process:
$X_t = Z_t \times X_{t-1}$ with $\mathrm{P}(Z_t = 1.03) = \mathrm{P}(Z_t = 1.02) = \mathrm{P}(Z_t = 0.96) = 1/3$.
For the stock price at time t, we can write:

$$X_t = X_0 \times \prod_{k=1}^{t} Z_k$$

and

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

Denote $\tilde{Z}_k = \log Z_k$, and observe that:

$$\begin{aligned} \mathrm{P}(\tilde{Z}_k = \log 1.03) &= \mathrm{P}(\tilde{Z}_t = \log 1.02) \\ &= \mathrm{P}(\tilde{Z}_t = \log 0.96) \\ &= \frac{1}{3} \end{aligned}$$

Define $\tilde{X}_t = \log X_t$, and then:

$$\tilde{X}_t = \tilde{X}_0 + \sum_{k=1}^{t} \tilde{Z}_k$$

Since the sample size $T = 260$ is sufficiently large, the trinomial process $\tilde{X}_t$ follows approximatively a normal distribution with following parameters:

$$\begin{aligned} \mu &= \mathsf{E}(\tilde{X}_t) \\ &= \mathsf{E}(\tilde{X}_0) + t \times \mathsf{E}(\tilde{Z}_1) \\ &= \log 100 + 260 \times \frac{1}{3}(\log 1.03 + \log 1.02 + \log 0.96) \\ &= 5.35 \\ \sigma^2 &= \mathsf{Var}(\tilde{X}_t) \\ &= \mathsf{Var}(\tilde{X}_0) + t \times \mathsf{Var}(\tilde{Z}_1) \end{aligned}$$

Since $\mathsf{Var}(\tilde{X}_0) = \mathsf{Var}(\log 100) = 0$ and $\mathsf{Var}(\tilde{Z}_1) = \mathsf{E}(\tilde{Z}_1^2) - \{\mathsf{E}(\tilde{Z}_1)\}^2$, see Härdle and Simar (2012), for the variance of $\tilde{X}_t$ holds:

$$\sigma^2 = t \times \left[\mathsf{E}(\tilde{Z}_1^2) - \{\ \mathsf{E}(\tilde{Z}_1)\}^2 \right]$$
$$= 260 \times \left[\frac{1}{3} \{(\log 1.03)^2 + (\log 1.02)^2 + (\log 0.96)^2\} \right]$$
$$- 260 \times \left\{ \frac{1}{9} (\log 1.03 + \log 1.02 + \log 0.96)^2 \right\} = 0.25$$

Since the trinomial process $\tilde{X}_t = \log X_t$ is approximately normally distributed, the stock price $X_t = \exp(\tilde{X}_t)$ is approximately lognormally distributed with mean:

$$\mathsf{E}(X_t) = \exp\left(\mu + \frac{1}{2}\sigma^2\right) = 238.89$$

and variance:

$$\mathsf{Var}(X_t) = \exp\left(\sigma^2 + 2 \times \mu\right) \times \left\{\exp\left(\sigma^2\right) - 1\right\} = 16355.48$$

For the 90 %-confidence interval, we therefore obtain:

$$X_{260} \in \left[\mathsf{E}(X_t) - 1.64 \times \sqrt{\mathsf{Var}(X_t)},\quad \mathsf{E}(X_t) + 1.64 \times \sqrt{\mathsf{Var}(X_t)} \right]$$

or

$$X_{260} \in [28.53, 449.25] ||$$

Exercise 7.3 (Binomial Tree). *A European put option with a maturity of 1 year and a strike price of 120 EUR is written on a non-dividend-paying stock. Assume the current stock price S_0 is 120 EUR, the volatility σ is 25 % per year, and the risk-free rate r is 6 % per year. Use a two-period binomial tree to value the option.*

(a) Construct an appropriate two-period pricing tree and check whether early exercise is optimal.

(b) Describe the replicating portfolio at each node. Verify that the associated trading strategy is self-financing, and that it replicates the payoff.

(a) We start with the calculation of the stock prices, in the two-period binomial tree $\triangle t = \frac{1}{2}$. The rate which the price moves up with equals:

$$u = \exp(\sigma + \sqrt{\triangle t}) = 1.19$$

The stock prices in the upper part of the tree are: $S_1^1 = S_0 \times u = 143.20$ and $S_2^2 = S_0 \times u^2 = 170.89$. The prices in the lower part of the tree move with the rate $d = 1/u$ such that: $S_1^0 = S_0/u = 100.56$ and $S_2^0 = S_0/u^2 = 84.26$.

After construction of the stock binomial tree, we can calculate the put option prices at maturity: $P(K = 120, S_2^j) = (K - S_3^j)^+$ for $j = 0, 1, 2$. To obtain option prices at period 1 and 0, we calculate the transition probability as follows:

$$(1-p)p = \frac{\sigma^2 \triangle t}{\{\log(u^2)\}^2} = 1, \quad \text{yielding} \quad p = \frac{1}{2}\,.$$

Applying the following equation, we calculate the put option prices at $t = 1$ and $t = 0$:

$$\mathrm{P}\left(K, S_n^k\right) = \exp(-r\triangle t)\left\{p\mathrm{P}\left(K, S_{j+1}^{k+1}\right) + (1-p)\mathrm{P}\left(K, S_{j+1}^{k}\right)\right\} \quad \text{for} \quad k, j = 0, 1\,.$$

Stock price	Option price		
170.89			0.00
143.21		0.00	
120.00	8.41		0.00
100.00		17.34	
84.26			35.74
Time	0.00	0.50	1.00

(b) The replicating portfolio at time 0 has $(0 - 35.74)/(170.89 - 84.26) = -0.41$ units of stock with value of -49.50, and long a bond with value $8.41 + 49.50 = 57.91$. The value of the replicating portfolio is equal to that of the option. This trading is self financed.

Exercise 7.4 (One Period Trinomial Model). *Show that the payoff of a call option cannot be replicated by stock and bond in a one period trinomial model. Assume zero interest rate for simplicity.*

Let S_0 be the price of a stock at time $t = 0$, and S^u, S^m,S^d be the corresponding upper, middle and down movement prices. Construct the replicating strategy of x stocks and y bonds for the call option with strike price K where $S^d < K < S^u$. At time $t = 1$ the strategy should produce the payoff:

$$\begin{cases} xS^u \ +y = (S^u - K)^+ \\ xS^m \ +y = (S^m - K)^+ \\ xS^d \ +y = (S^d - K)^+ \end{cases}$$

This system has a solution only when $S^u = S^m$ or $S^m = S^d$, which in fact reduces this model to the one period binomial model.

Exercise 7.5 (Hedging Strategy in One-Period Trinomial Model). *Find the hedging strategy for a call option in a one period trinomial model such that the quadratic hedging error is minimal. The quadratic hedging error is understood as the square distance between the actual payoff of the option and the final value of the hedge portfolio. Assume zero interest rate for simplicity.*

Consider the trinomial model as in Exercise 7.4. If $S^u > S^m > S^d$ and $S^d < K < S^u$ one cannot perfectly replicate the payoff. Hedging errors $\varepsilon^u, \varepsilon^m, \varepsilon^d$ appear in the system of equations.

$$\begin{aligned} xS^u + y &= (S^u - K)^+ + \varepsilon^u \\ xS^m + y &= (S^m - K)^+ + \varepsilon^m \\ xS^d + y &= (S^d - K)^+ + \varepsilon^d \end{aligned}$$

The hedging strategy $(\widehat{x}, \widehat{y})^\top$ minimizing the quadratic hedging error is given by the solution of the following least squares problem:

$$\min_{x,y} ||A\binom{x}{y} - b||^2$$

where

$$A = \begin{pmatrix} S^u & 1 \\ S^m & 1 \\ S^d & 1 \end{pmatrix}, b = \begin{pmatrix} (S^u - K)^+ \\ (S^m - K)^+ \\ (S^d - K)^+ \end{pmatrix}$$

The solution is:

$$(\widehat{x}, \widehat{y})^\top = (A^\top A)^{-1} A^\top b$$

Exercise 7.6 (Risk Neutral Probabilities). *Consider the one period trinomial model with the price $S_0 = 100$ at time $t = 0$. At time $t = 1$ the three possible stock prices are $S^u = 120$, $S^m = 100$ and $S^d = 80$. Find the risk neutral probabilities q^u, q^m, q^d of the* up, middle, *and* down *movements such that the price of the call option with strike K is equal to the price of the hedging portfolio minimizing the quadratic hedging error (see Exercise 7.5) at time $t = 0$. Consider the cases $K_1 = 110$, $K_2 = 100$ and $K_3 = 90$. Assume zero interest rate for simplicity.*

The risk neutral condition together with unit requirement for the sum of probabilities yield two equations. The third one comes from the call option pricing scheme i.e.

$$\begin{aligned} S^u q^u + S^m q^m + S_m q^d &= S_0 \\ q^u + q^m + q^d &= 1 \\ (S^u - K)^+ q^u + (S^m - K)^+ q^m + (S^u - K)^+ q^d &= \widehat{x} S_0 + \widehat{y} \end{aligned} \tag{7.1}$$

where the form of $(\widehat{x}, \widehat{y})^\top$ is given in Exercise 7.5. Solving the system (7.1) for K_1, K_2 and K_3 gives the probabilities $q^u = q^m = q^d = 1/3$. Note that the probabilities do not depend on K. Check also that for different price trees one obtains different risk neutral probabilities which are again independent on the choice of K.

SFStrinomialSB

Exercise 7.7 (Three-step implied Binomial Tree). *Construct a three-step implied binomial tree for stock prices, transition probabilities and Arrow-Debreu prices using the Derman-Kani algorithm. Assume the current value of the underlying stock* $S = 100$*, with no dividend, time to maturity* $T = 1$ *year and the annually compounded riskless interest rate* $r = 3\,\%$ *per year for all time expirations. In contrast to Cox-Ross-Rubinstein (CRR) binomial tree we use a nonconstant function for the implied volatility, let us assume the following convex function of moneyness, defined as* $\log(K/S_t)$*:*

$$\widehat{\sigma}(K, S_t) = -0.2/[\{\log(K/S_t)\}^2 + 1] + 0.3 \ .$$

First, we set the starting node of level zero to the current value of the underlying stock: $S_0^0 = 100$. In the next step, we calculate the stock price in the upper node of the first level S_1^1 from the equation:

$$\begin{aligned} S_1^1 &= \frac{S_0^0 \{C\,(S, \triangle t)\exp(r\triangle t) + \lambda_0^0 S_0^0 - \rho_u\}}{\lambda_0^0 F_0^0 - \exp(r\triangle t) C\left(S_0^0, \triangle t\right) + \rho_u} \\ &= 105.94\ , \end{aligned}$$

where $\lambda_0^0 = 1$, $\rho_u = \sum_{j=1}^{0} \lambda_0^j (F_0^j - S_0^0) = 0$ and $F_0^0 = \exp(r\triangle t) S_0^0 = 101.01$. Using the implied volatility for $K = S_0^0$ and $S_t = S$, $\sigma = \widehat{\sigma}(S_0^0, S)$, we calculate the call option price for strike price $K = S_0^0$ from the CRR binomial tree, $C\left(S_0^0, \triangle t\right) = 3.37$.

As we calculate the stock prices in an odd level, the stock price in the lower node must adjust the logarithmic spacing condition:

$$S_1^0 = \frac{\left(S_0^0\right)^2}{S_1^1} = 94.39.$$

Now, we can calculate the transition probability of making a transition from node (0,0) to node (1,1):

$$p_1^0 = \frac{F_0^0 - S_1^0}{S_1^1 - S_1^0} = 0.5726\ .$$

The Arrow-Debreu prices in the first level are:

$$\begin{aligned} \lambda_1^0 &= \exp(-r\triangle t)\left\{\lambda_0^0(1 - p_1^0)\right\} = 0.5669 \\ \lambda_1^1 &= \exp(-r\triangle t)\left(\lambda_0^0 p_1^0\right) = 0.4231 \end{aligned}$$

In the next (even) level, we start with the central node and define $S_2^1 = S_0^0 = 100$. Then we use the formula for stock price in the upper node:

$$S_2^2 = \frac{S_2^1 \left\{C\left(S_1^1, 2\triangle t\right)\exp(r\triangle t) - \rho_u\right\} - \lambda_1^1 S_1^1 \left(F_1^1 - S_2^1\right)}{C\left(S_1^1, 2\triangle t\right)\exp(r\triangle t) - \rho_u - \lambda_1^1\left(F_1^1 - S_2^1\right)} = 112.38\,,$$

with $F_1^1 = \exp(r\triangle t)S_1^1 = 107.00$ and $\rho_u = \sum_{j=2}^{1} \lambda_1^j (F_1^j - S_1^1) = 0$. We obtain the call option price from the CRR binomial tree with $\sigma = \widehat{\sigma}(S_1^1, S)$, $C\left(S_1^1, 2\triangle t\right) = 2.05$.

To compute the stock price in the lower node we use the following formula:

$$S_2^0 = \frac{S_2^1 \left\{\exp(r\triangle t)\mathrm{P}\left(S_1^0, 2\triangle t\right) - \rho_l\right\} - \lambda_1^0 S_1^0 \left(F_1^0 - S_2^1\right)}{\exp(r\triangle t)\mathrm{P}\left(S_1^0, 2\triangle t\right) - \rho_l + \lambda_1^0\left(F_1^0 - S_2^1\right)}\,,$$

with $F_1^0 = \exp(r\triangle t)S_1^0 = 95.34$ and $\rho_l = \sum_{j=0}^{-1} \lambda_1^j \left(S_1^0 - F_1^j\right) = 0$. The put option price is calculated from the CRR binomial tree with $\sigma = \widehat{\sigma}(S_1^0, S)$, $P\left(S_1^0, 2\triangle t\right) = 0.97$.

The transition probabilities from node (1,0) to node (2,1) and from node (1,1) to node (2,2) are:

$$p_1^1 = \frac{F_1^0 - S_2^0}{S_2^1 - S_2^0} = 0.5658\,,$$

$$p_2^1 = \frac{F_1^1 - S_2^1}{S_2^2 - S_2^1} = 0.5807\,.$$

Now, we can calculate also the Arrow-Debreu prices in the second level:

$$\lambda_2^0 = \exp(-r\,\triangle t)\lambda_1^0(1 - p_1^1) = 0.3176$$

$$\lambda_2^1 = \exp(-r\,\triangle t)\{\lambda_1^0 p_1^1 + \lambda_1^1(1 - p_2^1)\} = 0.4870$$

$$\lambda_2^2 = \exp(-r\,\triangle t)\lambda_1^1 p_2^1 = 0.1757\,.$$

Analogously, we proceed the calculation for the last level n = 3 to get the complete IBT. Check your results with the quantlet http://www.quantlet.com/mdstat/codes/sfs/SFSIBTdk.html `SFSIBTdk`, see Figs. 7.1–7.3.

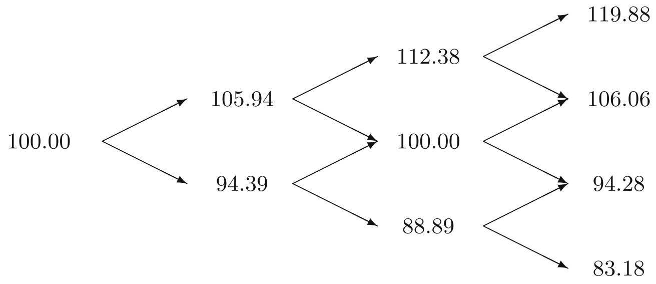
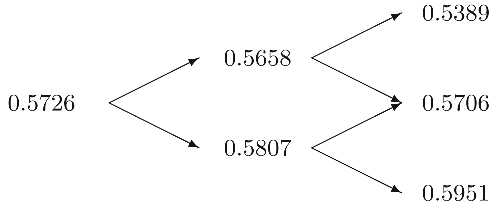


Fig. 7.1 DK stock price tree

Fig. 7.2 DK transition probability tree



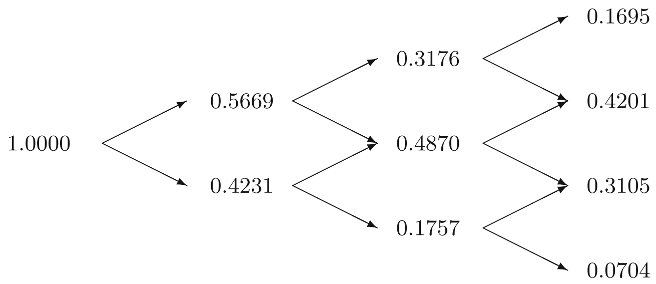


Fig. 7.3 DK Arrow-Debreu price tree

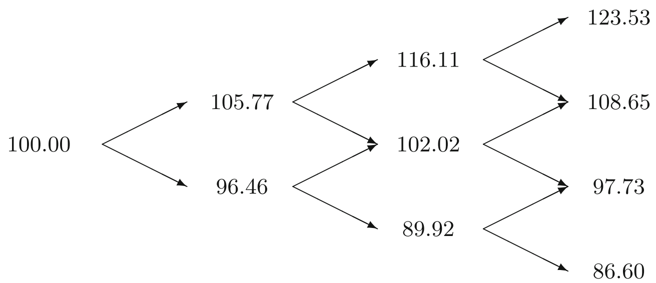


Fig. 7.4 BC stock price tree

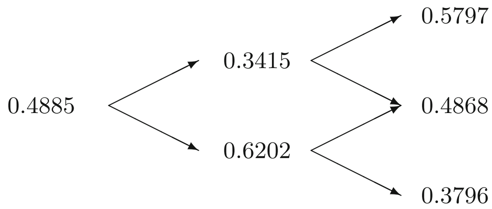


Fig. 7.5 BC transition probability tree

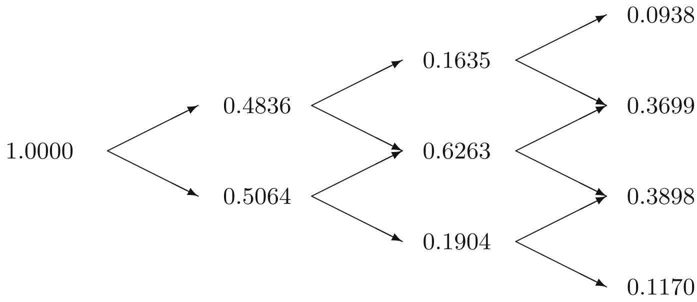


Fig. 7.6 BC Arrow-Debreu price tree

Fig. 7.7 Arrow-Debreu prices from the BC tree

Exercise 7.8 (Method of Barle-Cakici). *Consider the call option from Exercise 7.7 and construct the IBTs using the method of Barle-Cakici (BC). Assume an exercise price* $K = 100$ *EUR/USD and compute the option price implied by the binomial tree. Make a plot of the Arrow-Debreu prices.*

First, we construct the BC IBT using the quantlet http://www.quantlet.com/mdstat/codes/sfs/SFSIBTbc.html `SFSIBTbc`. The BC construction is similar to the DK algorithm from Exercise 7.8, one set the central nodes $S^i_{n+1} = S^0_0 \exp(rn\triangle t)$ and uses the Black-Scholes call and put option prices $C(F^i_n, (n+1)\triangle t)$ and $P(F^i_n, (n+1)\triangle t)$, respectively.

To compute the call option price from an IBT, we use the Arrow-Debreu prices and the stock prices in the last level of the tree. In our discrete model, the call option price is (Figs. 7.4–7.7):

$$C(K, (n+1)\triangle t) = \sum_{i=0}^{n+1} \lambda^i_{n+1} \max(S^i_{n+1} - K, 0)\ . \tag{7.2}$$

The stock prices in the last level are: $S^0_3 = 86.60,\ S^1_3 = 97.73,\ S^2_3 = 108.65,\ S^3_3 = 123.53$. Corresponding Arrow-Debreu prices in the third level are $\lambda^0_3 = 0.1170,\ \lambda^1_3 = 0.3898,\ \lambda^2_3 = 0.3699,\ \lambda^3_3 = 0.0938$. For the call price at maturity and exercise price $K = 100$ then follows:

$$C(100, 1) = \sum_{i=0}^{3} \lambda^i_3 \max(S^i_3 - K, 0) = 5.41 \quad \text{EUR/USD}\ .$$

Chapter 8
American Options

Opsi Amerika
Memang di dalam kehidupan ini tidak ada yang pasti. Tetapi kita harus berani memastikan apa-apa yang ingin kita raih.
Indeed, there is uncertainty in this life. But we must dare to make sure what we want to achieve.

Up to now we have considered mainly European options. This chapter however focuses on American Options. An American option is an option that can be exercised anytime during its life. The time at which the holder chooses to exercise the options depends on the spot price of the underlying asset S_t. In this sense, the exercising time is a random variable itself. It is obvious that the Black-Scholes differential equations still hold as long as the options are not exercised. However the boundary conditions are so complicated that an analytical solution is not possible.

The right to early exercise implies that the value of an American option can never drop below its intrinsic value. For example, the value of an American put should not go below $\max(K - S_t, 0)$ with the exercise price K. In contrast, this condition does not hold for European options. That is because American puts would be exercised before expiry date if the value of the option drop below the intrinsic value. Because of their freedom to exercise American options at any point during the life of the contract, they are more valuable than European options which can only be exercised at maturity.

Exercise 8.1 (Relations). *Explain the relation between American call and put, and the following: value of the underlying asset, exercise price, stock volatility, interest rate, time to exercise date.*

The payoff of an American call option with exercise time $t^* \in [t_0, T]$ is $\max(S_{t^*} - K, 0)$, and is $\max(K - S_{t^*}, 0)$ for the put option. Increasing the value of the underlying asset would possibly increase the call option price. While for the put option, the price will possibly decrease.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_8,

In the same way, increasing the exercise price K would possibly decrease the payoff $\max(S_{t^*} - K, 0)$ for call options and raise the payoff $\max(K - S_{t^*}, 0)$ for put options.

Increasing stock volatility and time to exercise date would both raise the risk of buying put and call, therefore increase the price.

Increasing the interest rate would increase the profit of saving money, so the price of a call would be higher, since the option allows one to save money before exercise. Since for put options one requires to buy the underlying in advance to sell at t^*, one loses the opportunity to save money in bank, and the cost of carry increases. Therefore the profit decreases, leading to a decrease of the put option prices.

The next table summarizes that how the American call and put options prices change when the corresponding variables change (increase(+)/decrease(−)):

Increase	Call option	Put option
Value of the underlying asset	+	−
Exercise price	−	+
Stock volatility	+	+
Interest rate	+	−
Time to exercise date	+	+

Exercise 8.2 (Price of an American Call Option). *Consider a stock whose price increases by* 20 % *or decreases by* 10 % *in each period. We are now at* $t = 0$ *and we have an American call option on this stock with an exercise price of* 10.5 *and a terminal value at* $T = 2$*. What will be the price of this American call option at* $t = 0$*? Will it be different from the price of a European call option? (Set the interest rate and dividend equal to* 0*, and denote the intrinsic value at time* T *as* C_T*.)*

$t = 0$	$t = 1$	$t = 2$	C_T
		14.4	3.9
	12		
10		10.8	0.3
	9		
		8.1	0.0

The movement of the stock prices can be seen above. Let us find the value of this call option at $t = 1$ on the upper branch of the tree. We will use the martingale measure approach to find the price of the call option.

$t = 1$	$t = 2$	C_T
	14.4	3.9
12		
	10.8	0.3

$$\begin{aligned} q \cdot 14.4 + (1-q) \cdot 10.8 &= 12 \\ q &= 1/3 \\ C_1 &= 1/3 \cdot 3.9 + 2/3 \cdot 0.3 \\ &= 1.5 \end{aligned}$$

The intrinsic value of the option at $t = 1$ is also $12 - 10.5 = 1.5$. So, the value of the American option will not change whether it will be exercised at $t = 1$ or kept until $t = 2$.

Now, let's look at the other branch:

$t = 1$	$t = 2$	C_T
	10.8	0.3
9		
	8.1	0.0

$$\begin{aligned} q \cdot 10.8 + (1-q) \cdot 8.1 &= 9 \\ q &= 1/3 \\ C_1 &= 1/3 \cdot 0.3 + 2/3 \cdot 0 \\ &= 0.1 \end{aligned}$$

The intrinsic value of the option at $t = 1$ is 0. So, to exercise the American call option, it will be optimal to wait until $t = 2$.

Now, we can calculate the value of this American call option at $t = 0$:

$t = 0$	$t = 1$	C_1
	12	1.5
10		
	9	0.1

$$\begin{aligned} q \cdot 12 + (1-q) \cdot 9 &= 10 \\ q &= 1/3 \\ C_0 &= 1/3 \cdot 1.5 + 2/3 \cdot 0.1 \\ &= 0.566 \end{aligned}$$

As we have seen, it is optimal to keep the American call option until the maturity. The early exercise possibility of the American call option does not create any value here. So it will not be early exercised and its price is equal to the European call option.

Exercise 8.3 (Option Value and Put-Call Parity). *A stock price is currently* 50. *The price can increase by a factor of* 1.10*, or fall by a factor of* 0.90. *The stock pays no dividends and the yearly discrete compounding interest rate is* 0.05. *Consider American put and call options on this stock with strike price* 50*, and 2 years' time to maturity and 1 year's step length.*

(a) What will be the price of this American call option at $t = 0$*? Will it be different from the price of a European call option?*
(b) What will be the price of this American put option with the same strike price?
(c) Does the put call parity hold?

(a) The stock price is:

$t = 0$	$t = 1$	$t = 2$
		60.5
	55	
50		49.5
	45	
		40.5

Risk neutral probability is

$$q = \frac{1 \cdot 1.05 - 0.9}{1.1 - 0.9} = 0.75$$

The call pricing tree without early exercise:

$t = 0$	$t = 1$	$t = 2$
		10.50
	7.50	
5.36		
	0.00	
		0.00

Note for the call, there are no nodes at which early exercise is optimal.

(b) The put pricing tree without early exercise:

$t = 0$	$t = 1$	$t = 2$
		0.00
	0.12	
0.71		0.50
	2.62	
		9.50

For the American put, early exercise is optimal at $t = 1$, $S_t = 45$ since the intrinsic value is $50 - 45 = 5$, which is larger than the value 2.62 of the option.

The American put pricing tree with early exercise:

$t = 0$	$t = 1$	$t = 2$
		0.00
	0.00	
1.19		0.00
	5.00	
		9.50

(c) Put call parity: $P = C - S + K/(1+r)^{\tau}$
European Options: $0.71 = 5.36 - 50 + 50/1.05^2$
American Options: $1.19 > 5.36 - 50 + 50/1.05^2$,
where $K/(1+r)^{\tau}$ is the discounted strike price.

For non-dividend paying stocks, the European and American calls are of the same value, but the American put is worth more than the European put. Since put-call parity holds for the European options, American puts are generally worth more than their simple parity values.

Exercise 8.4 (Option Value and Put-Call Parity). *Consider the same model as in Exercise 8.3. Assume that there is no dividend, and use continuous compounding interest rate.*

(a) Find the value of both European and American call options with strike prices of 50 and maturities of 2 years. The yearly compounding risk-free rate is 5 %.
(b) Find the value of both European and American put options with strike prices of 50 and maturities of 2 years. The yearly compounding risk-free rate is 5 %.
(c) Is the put-call parity relation satisfied by the European options? For the American ones? Would you predict that the American put price will be higher than its parity value in general? Explain.

(a) Again we would first calculate the price movement and intrinsic value C_T accordingly.

$t = 0$	$t = 1$	$t = 2$	C_T
		60.5	10.5
	55		
50		49.5	0.0
	45		
		40.5	0.0

Based on martingale approach, the call option price at time t is just the discounted expected return.

So at $t = 1$, we have:
When $S_t = 55$, the call option price is

$$\{q \cdot 10.5 + (1-q) \cdot 0.00\} \times \exp(-r) = 7.554,$$

where q is the risk neutral probability calculated from:

$$\begin{aligned}\{q \cdot 60.5 + (1-q) \cdot 49.5\} \times \exp(-r) &= 55 \\ q &= 0.756\end{aligned}$$

When $S_t = 45$, similarly, we get that the call price is 0.
Comparing the price with the intrinsic value at t, we notice:

$$7.554 \geq (55 - 50)$$

Therefore the American call option would not be early exercised, and the European and American call options would be of the same price C_0.

$$\begin{aligned}\{q \cdot 55 + (1-q) \cdot 45\} \times \exp(-r) &= 50.0 \\ q &= 0.756 \\ \{0.76 \cdot 7.55 + (1-0.76) \cdot 0\} \times \exp(-r) &= C_0 \\ &= 5.435\end{aligned}$$

(b) The stock prices and intrinsic values are shown as below:

$t = 0$	$t = 1$	$t = 2$	C_T
		60.5	0.0
	55		
50		49.5	0.5
	45		
		40.5	9.5

At $t = 1$, we have:
When $S_t = 55$, the put option price with $q = 0.756$ is,

$$\{q \cdot 0.0 + (1-q) \cdot 0.5\} \times \exp(-r) = 0.116$$

When $S_t = 45$, it is,

$$\{q \cdot 0.5 + (1-q) \cdot 9.5\} \times \exp(-r) = 2.561$$

As $2.561 \leq (50 - 45)$, the American put option could be early exercised, and American and European option prices deviate.

Then the European put option price would be:

$$C_0 = \{0.756 \cdot 0.116 + (1 - 0.756) \cdot 2.561\} \times \exp(-r)$$
$$= 0.677,$$

whereas the American put option price is:

$$C_0 = \{0.756 \cdot 0.116 + (1 - 0.756) \cdot 5\} \times \exp(-r)$$
$$= 1.240.$$

(c) Yes, plug in the European option price derived above to check whether:

$$C - S + K \times \exp(-r\tau) = P,$$

with $\tau = 2$.

Since we have:

$$\begin{aligned} C - S + K \times \exp(-r\tau) &= 5.435 - 50 + 50 \cdot 0.904 \\ &= 0.677 \\ &= P \end{aligned}$$

The put call parity is satisfied for European option prices.

But with $P = 1.240$, we know $P > C - S + K \times \exp(-r\tau)$ for American option prices, thus the put call parity is not satisfied. That is due to the early exercise of the put option.

Exercise 8.5 (Option Value, Dividends and Put-Call Parity). *Consider the same model as in Exercise 8.3. However, this time we know that there will be a dividend payment at $t = 1$ equal to 5.*

(a) Find the value of both European and American call options with strike prices of 50 and maturities of 2 years. The yearly compounding risk-free rate is 5 %.
(b) Find the value of both European and American put options with strike prices of 50 and maturities of 2 years. The yearly compounding risk-free rate is 5 %.
(c) Is the put-call parity relation satisfied by the European options? For the American ones? Would you predict that the American put price will be higher than its parity value in general? Explain.

(a) We first calculate the price movement and intrinsic value C_T using a discounted initial value.

$$\tilde{S}_0 = S_0 - D_1 \times \exp(-r)$$

$t=0$	$t=1$	$t=2$	C_T
		54.7	4.7
	49.8		
45.2		44.7	0.0
	40.7		
		36.6	0.0

Then we calculate the European call value:

$t=0$	$t=1$	$t=2$
		4.75
	3.41	
2.46		0.00
	0.00	
		0.00

But for the American call price, we need to compare the value with intrinsic value $\max(S_t - K + D, 0)$ at each state t. Then we have the value as below:

$t=0$	$t=1$	$t=2$
		4.75
	4.77	
3.43		0.00
	0.00	
		0.00

So we see that the American call option with dividend would possibly be early exercised.

(b) Following the price movement in (a), we have the European put option value:

$t=0$	$t=1$	$t=2$
		0.00
	1.21	
2.45		5.21
	4.28	
		13.35

Compared with the intrinsic value $\max(K - S_t - D, 0)$, the American put option has the same value. So we notice that dividend would decrease the chance of early exercise for American put.

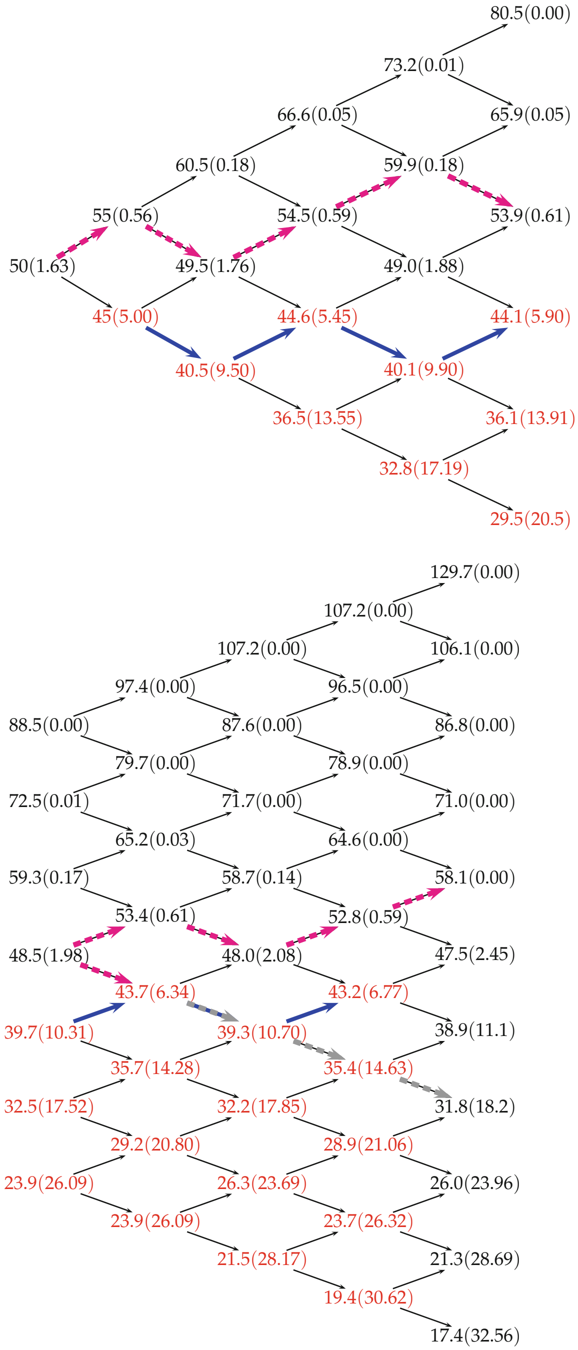


Fig. 8.1 Binomial tree for stock price movement and option value (in parenthesis)

(c) Put call parity in this setting:

$$C - S + K \times \exp(-r\tau) + D = P,$$

Since we have:

$$\begin{aligned} C - S + K \times \exp(-r\tau) + D &= 2.46 - 50 + 50 \cdot 0.904 + 5 \cdot 0.951 \\ &= 2.45 \\ &= P \end{aligned}$$

The put call parity is satisfied for European option.

However, for American option, we notice a higher price for American call option due to early exercise. So

$$C - S + K \times \exp(-r\tau) + D > P.$$

Exercise 8.6 (Option Value in 10-step Binomial Model). *Consider the same model as in Exercise 8.3, but extend the binomial model to 10 steps in 2 years for an American put option.*

The movement of the stock price and American call value (in parenthesis) are shown in Fig. 8.1. The critical nodes (marked red) show the nodes where the value is less than the intrinsic value. Once touching those nodes, the option would be early exercised. Thus, the blue line is the critical bound for stock prices. The put would be early exercised if the stock price touches the bound. There are two paths of stock price (in magenta) demonstrating the above idea. The two paths branch at the 6th step when the stock price is 48.5. For one path, we see that when the stock price falls to 43.7 in the 7th step, the put would be early exercised. While for the other, the put would not be early exercised because it did not touch the critical bound, so the vanished path after is shown in gray. Check your results with the quantlet SFSbitreeNDiv.

Chapter 9
Exotic Options

A man with one watch always knows what time it is, a man with two watches is never sure.

Exotic options are financial derivatives which are more complex than normally traded options (vanilla options). They are mainly used in OTC-trading (over the counter) to meet special needs of corporate customers. For example, a compound option allows one to acquire an ordinary option at a later date, and a chooser option is a form of the compound option where the buyer can decide at a later date which type of option he would like to have.

Compared to straight call and put options, exotic options are more difficult to price. However, we can still obtain some insights by using a standard approach, such as the Black Scholes formula or binomial trees to valuate them. But indeed, exotic options may lead to challenging problems in valuation and hedging.

Exercise 9.1 (Compound Option). *A compound option is also called option on option. It allows the purchaser to acquire an ordinary option at a later date. Consider a European Call-on-a-Call option, with the first expiration date T_1, the second expiration date T_2, the first strike price K_1, and the second strike price K_2. (a) Determine the value of the compound option at time T_1. (b) Let $T_1 = 4$ months, $T_2 = 12$ months, $K_1 = 25$, $K_2 = 220$, initial value of the asset, volatility $\sigma = 0.23$, $r = 0.034$ and $S_{T_1} = 230$, calculate the value of compound option at time T_1.*

(a) The purchaser of a compound option has the right to buy a new call option at T_1 for the price K_1, and the new call has maturity T_2 and strike price K_2. So whether the purchaser buys the call would depend on whether the value of the call is higher than K_1 at T_1. At time T_1, we have the value of the compound option:

$$C^{Compound} = \max\{0, C_{BS}(S_{T_1}, T_2 - T_1, K_2) - K_1\}, \tag{9.1}$$

where $C_{BS}(S_{T_1}, T_2 - T_1, K_2)$ is the Black Scholes call option price.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_9,

(b) $C_{BS}(S_{T_1}, T_2 - T_1, K_2) = 25.1614$, and $C^{Compound} = \max\{0, C_{BS}(S_{T_1}, T_2 - T_1, K_2) - K_1\} = 0.1614$

Exercise 9.2 (Chooser Option). *A chooser option (preference option) is a path dependent option for which the purchaser pays an up-front premium and has the choice of exercising a vanilla put or call on a given underlying at maturity. The purchaser has a fixed period of time to make the choice. At time $0 < T_0 < T$, the purchaser chooses the option with the higher value.*

(a) Give the payoff of the chooser option.
(b) In a non-arbitrage framework with $S_t^0 = (1+r)^t (t = 0, \cdots, T, r > -1)$, show that the price of the call and the put option is equivalent to the price of a chooser option:

$$C^{Chooser} = C^{Call} + C^{Put} \tag{9.2}$$

where C^{Call} denotes the call option with underlying price S_t, strike price K and maturity T. C^{Put} defines the put option with strike price $K(1+r)^{T_0-T}$ and maturity T_0. (Hint: $\{(1+r)^{-t}S_t\}$ is a martingale.)

(a) Let $(S_t)_{0\le t\le T}$ be the stock price process.
The payoff of a chooser option is therefore:

$$\begin{aligned}
&\text{Payoff}\\
&= \max\{(S_T - K), 0\}\mathbf{1}(C_{T_0}^{Call} \ge C_{T_0}^{Put}) + \max\{(K - S_T), 0\}\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})\\
&= \max\{(S_T - K), 0\}\mathbf{1}(C_{T_0}^{Call} \ge C_{T_0}^{Put}) + \max\{(S_T - K), 0\}\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})\\
&\quad + \max\{(K - S_T), 0\}\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put}) - \max\{(S_T - K), 0\}\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})\\
&= \max\{(S_T - K), 0\} + (K - S_T)\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})
\end{aligned}$$

(b) Following the law of iterated expectations and the martingale property, the value of the chooser option is the discounted expected payoff:

$$\begin{aligned}
&C^{Chooser}\\
&= \mathsf{E}^Q[\max\{(S_T - K), 0\}/S_T^0] + \mathsf{E}^Q[\{(K - S_T)/S_T^0\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})]\\
&= \mathsf{E}^Q[\max\{(S_T - K), 0\}/S_T^0] + \mathsf{E}^Q[\mathsf{E}^Q[\{K - S_T\}/S_T^0|\mathcal{F}_{T_0}]\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})]\\
&= \mathsf{E}^Q[\max\{(S_T - K), 0\}/S_T^0] + \mathsf{E}^Q[\{K - S_{T_0}(1+r)^{T-T_0}\}/S_T^0\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})]\\
&= \mathsf{E}^Q[\max\{(S_T - K), 0\}/S_T^0] + \mathsf{E}^Q[\{K(1+r)^{T_0-T} - S_{T_0}\}/S_{T_0}^0\\
&\qquad \times \mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put})]
\end{aligned} \tag{9.3}$$

Using the Call-Put Parity:

$$C_{T_0}^{Call} - C_{T_0}^{Put} = S_{T_0} - K(1+r)^{T_0-T}$$

One can show that:

$$\mathbf{1}(C_{T_0}^{Call} < C_{T_0}^{Put}) = \mathbf{1}\{S_{T_0} < K(1+r)^{T_0-T}\}$$

The Eq. (9.3) can be rewritten as follows:

$$\begin{aligned} &\mathsf{E}^Q[\max\{(S_T - K), 0\}/S_T^0] + \mathsf{E}^Q[\{K(1+r)^{T_0-T} - S_{T_0}\} \\ &\times \mathbf{1}(S_{T_0} < K(1+r)^{T_0-T})/S_{T_0}^0] \\ = &\ \mathsf{E}^Q[\max\{(S_T - K), 0\}/S_T^0] + \mathsf{E}^Q[\max\{K(1+r)^{T_0-T} - S_{T_0}, 0\}/S_{T_0}^0] \end{aligned}$$

The first expectation is the call price with strike price K and maturity T, and the second expectation is the price of the put option with strike price $K(1+r)^{T_0-T}$ and time to maturity T_0. Therefore, it holds:

$$C^{Chooser} = C^{Put} + C^{Call}$$

Exercise 9.3 (Cliquet Option). *In a cliquet option, the strike price periodically resets before the expiration date is reached. If the expiration date is reached, and the underlying price is below the strike price, the option will expire worthless, and the strike will be reset to the lowest value of the underlying price. If the expiration date is reached and the underlying value is higher than the strike price, the purchaser of the option will earn the difference and the strike price will reset to the higher underlying price. Consider a cliquet call with maturity* $T = 3$ *years and strike price* $K_1 = 100$ *in the first year. Suppose that the underlying in the 3 years are* $S_1 = 90$, $S_2 = 120$, $S_3 = 130$. *What is the payoff of the cliquet option?*

The payoff of the cliquet call at maturity T:

$$\max\{(S_{t_1}, S_{t_2}, \ldots, S_{t_n=T}) - S_{t_1}\}$$

In the first year, the underlying was $S_1 = 90$, the cliquet option would expire worthless. The new strike price for the second year will be set to $K_2 = 90$. In the second year $S_2 = 120$, then the contract holder will receive a payoff 30 and the strike price would reset to the new level of $K_3 = 120$. At maturity, the payoff would be $S_3 - K_3 = 10$. The total payoff would be $30 + 10 = 40$.

Also we could use:

$$\max\{(S_{t_1}, S_{t_2}, S_{t_3}) - S_{t_1}, 0\} = 130 - 90 = 40$$

In sum, the payoff is 40 in this example.

Exercise 9.4 (Barrier Option). *A barrier option changes its value in leaps as soon as the stock price reaches a given barrier, which can also be time dependent. A European down-and-in call is a barrier option which starts to be active only when*

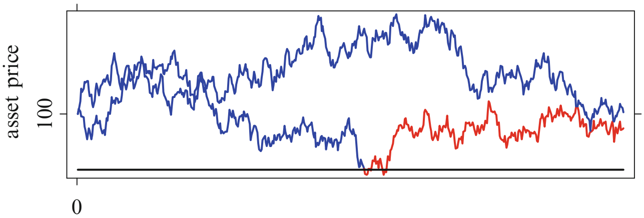


Fig. 9.1 Two possible paths of the asset price. When the price hits the barrier (*lower path*), the option expires worthless. SFSrndbarrier

the underlying $S_t \leq B$ at any any time $0 \leq t \leq T$, and a European down-and-out call is a barrier option which expires worthless as long as the underlying $S_t \geq B$ at any time $0 \leq t \leq T$. The two options share the same maturity time T and strike price K. Explain why the down-and-in and the down-and-out call together have the same effect as a normal European call (In-Out-Parity).

We know that at a time $0 \leq t \leq T$, either $S_t \leq B$ or $S_t \geq B$ happens. It follows that one and only one of the two options would be valid at maturity T. Thus the down-and-in and the down-and-out call have the same payoff with a normal call at T. By the no-arbitrage principle, we know that the down-and-in and the down-and-out call have the same price with the normal call at t. See Fig. 9.1 for an example of knock-out option.

Exercise 9.5 (Forward Start Option). *Forward start options are options whose strike is unknown at the beginning, but will be determined at an intermediate time t. So a forward start option is similar to a vanilla option except for not knowing the strike price at the moment of purchase. The strike is usually determined by the underlying price at time t. Let S_t denote a random path in a binomial tree. Let $a > 0$ and $b < 0$ denote the upward rate and downward rate. r is the risk free interest rate.*

Let $0 < t < T$. Calculate the price and a Delta hedge for the forward start option, whose payoff is given by

$$\max\{(S_T/S_t - K), 0\}$$

where K is the strike price and S_T is the stock price at maturity T.

The discounted payoff of the forward start option is

$$(1+r)^{-T} \max\{(S_T/S_t - K), 0\},$$

and can also be written as

$$(1+r)^{-T} \max[\{(1+a)^s(1+b)^{T-t-s} - K\}, 0],$$

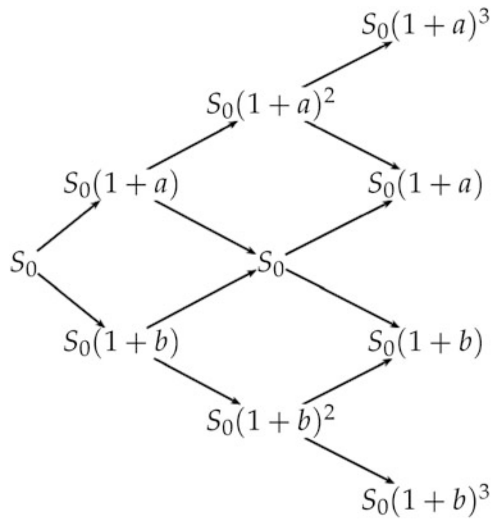


Fig. 9.2 Binomial tree for stock price movement at time $T = 3$

where $a > 0$ and $b < 0$ are the rates of going up and down. Also we assume that $1 + a = 1/(1 + b)$ for the recombining property.

As shown in Fig. 9.2, S_t can be expressed as the initial price adjusted by $0 \le k \le t$ upward movements rate and $t - k$ downward movements rate. S_T can be expressed as the initial price adjusted by $0 \le k + s \le T$ upward movements rate and $T - k - s$ downward movements rate.

$$S_t = S_0(1+a)^k(1+b)^{t-k}$$
$$S_T = S_0(1+a)^{k+s}(1+b)^{T-k-s}$$

Under the risk neutral probability measure, the price of the forward start option (C^F) equals to the expected discounted payoff:

$$\begin{aligned} C^F &= (1+r)^{-T}\mathsf{E}^Q\left[\max\{(S_T/S_t - K)\,,0\}\right] \\ &= (1+r)^{-T}\sum_{s=0}^{T-t}\max[\{(1+a)^s(1+b)^{T-t-s} - K\},0]\binom{T-t}{s}q^s(1-q)^{T-t-s} \end{aligned}$$

where $q = (r - b)/(a - b)$ is the risk neutral probability of upward movements.

For a hedging strategy, suppose we know the stock price up to time l ($l > t$), then the option value at time l can still be calculated:

$$\begin{aligned} v_l(S_0,\ldots,S_l) = (1+r)^{l-T}\sum_{s=0}^{T-l}\max\left[\left\{\frac{S_l}{S_t}(1+a)^s(1+b)^{T-l-s} - K\right\},0\right] \\ \times\binom{T-l}{s}q^s(1-q)^{T-l-s} \end{aligned}$$

Then we can approximate Delta at l by

$$\begin{aligned}\Delta_l &= \left[\frac{v_t\{S_0,\dots,S_{l-1},S_{l-1}(1+a)\} - v_t\{S_0,\dots,S_{l-1},S_{l-1}(1+b)\}}{S_{l-1}(a-b)}\right](1+r)^l \\ &= \binom{T-t}{s} q^s(1-q)^{T-t-s} \\ &\quad \times \sum_{s=0}^{T-l} \max\left[\left\{\frac{S_{l-1}(1+a)}{S_t}(1+a)^s(1+b)^{T-l-s} - K\right\},0\right] \\ &\quad - \left[\max\left\{\frac{S_{l-1}(1+b)}{S_t}(1+b)^s(1+a)^{T-l-s} - K\right\},0\right] \\ &\quad \times \frac{1}{(a-b)S_{l-1}}\end{aligned}$$

Thus we can buy $-\Delta_l$ stocks at time l to hedge.

When $l \leq t$, $\Delta_l = 0$.

Exercise 9.6 (Forward Start Option). *Consider a call option with forward start $t = 4$ months from today ($t = 0$). The option starts at $K = 1.1S_t$, time to maturity is $T = 1$ year from today, the initial stock price S_0 is 60, the risk free interest rate is $r = 9\,\%$, the continuous dividend yield is $d = 3\,\%$, and the expected volatility of the stock is $\sigma = 20\,\%$. What is the price of this forward start option?*

The value of a forward start option C^F is given by:

$$\begin{aligned}C^F = S_0 \exp\{(b-r)t\}\{&\exp\{(b-r)(T-t)\}\Phi(d_1) \\ &-K/S_t \exp\{-r(T-t)\}\Phi(d_2)\}\end{aligned} \tag{9.4}$$

where

$$d_1 = \frac{\log(S_t/K) + (b+\sigma^2/2)(T-t)}{\sigma\sqrt{(T-t)}}$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)}$$

and $b = r - d$ as the cost of carry.

First, we calculate accordingly:

$$d_1 = -0.2571$$

$$d_2 = -0.4204$$

and then:

$$\Phi(d_1) = 0.3985$$

$$\Phi(d_1) = 0.3371$$

Plug in (9.4), the forward option price is:

$$C^F = 2.4589$$

Exercise 9.7 (Product Call Option). *Find a formula for a European-style "product call" with payoff* $\max(S_t^1 S_t^2 - K, 0)$*, where* S_t^1 *and* S_t^2 *are the prices of assets with correlated random increments. Apply the Black-Scholes assumptions, i.e. suppose that* S_t *follows a geometric Brownian motion.*

If W_t^1 and W_t^2 are two independent Brownian motions, then

$$B_t^1 = W_t^1$$
$$B_t^2 = \rho W_t^1 + \sqrt{1-\rho^2}W_t^2$$

define two correlated Brownian motions with correlation $\rho \in [-1, 1]$.

The stock price processes can be written as:

$$S_t^1 = S_0^1 \exp\left\{\left(\mu_1 - \frac{\sigma_1^2}{2}\right)t + \sigma_1 B_t^1\right\}$$
$$S_t^2 = S_0^2 \exp\left\{\left(\mu_2 - \frac{\sigma_2^2}{2}\right)t + \sigma_2 B_t^2\right\}$$

leading to

$$S_t^1 = S_0^1 \exp\left\{\left(\mu_1 - \frac{\sigma_1^2}{2}\right)t + \sigma_1 W_t^1\right\}$$
$$S_t^2 = S_0^2 \exp\left\{\left(\mu_2 - \frac{\sigma_2^2}{2}\right)t + \sigma_2\left(\rho W_t^1 + \sqrt{1-\rho^2}W_t^2\right)\right\}$$

Now the Girsanov transformation ξ_T (Theorem 22.4 in Franke et al. (2011)) for the two processes of discounted stock prices $\tilde{S}_t^1 = \exp(-rt)S_t^1$ and $\tilde{S}_t^2 = \exp(-rt)S_t^2$ has to be found. Some thoughts about the binomial formula suggest:

$$\frac{dQ}{dP} = \xi_T = \exp\left(\theta_1 W_T^1 + \theta_2 W_T^2 - \frac{1}{2}\theta_1^2 T - \frac{1}{2}\theta_2^2 T\right) \tag{9.5}$$

with θ_1 and θ_2 the unknown market price of risk (MPR). Recall that ξ_t has to fulfill the property of $\tilde{S}_t^1 = \exp(-rt)S_t^1$ and $\tilde{S}_t^2 = \exp(-rt)S_t^2$ being Q-martingales. In terms of (9.5) this means that the equations:

$$\mathsf{E}^Q[\tilde{S}_t^j|\mathcal{F}_s] = \tilde{S}_s^j,\ j = 1, 2$$

have to hold for all $s \le t$, where

$$\mathsf{E}^Q[\tilde{S}_t|\mathcal{F}_s] = \mathsf{E}\left[\frac{dQ}{dP}\tilde{S}_t|\mathcal{F}_s\right] = \mathsf{E}[\xi_t\,\tilde{S}_t|\mathcal{F}_s].$$

Using (9.5) and the interest rate r we obtain for the parameters θ_1 and θ_2:

$$\theta_1 = \frac{r-\mu_1}{\sigma_1} \tag{9.6}$$

$$\theta_2 = \frac{1}{\sqrt{1-\rho^2}}\left(\frac{r-\mu_2}{\sigma_2} - \rho\frac{r-\mu_1}{\sigma_1}\right) \tag{9.7}$$

Thus the Q-Brownian motions are:

$$W_t^{*1} = W_t^1 - \theta_1 t = W_t^1 - \frac{r-\mu_1}{\sigma_1}t$$

$$W_t^{*2} = W_t^2 - \theta_2 t = W_t^2 - \frac{1}{\sqrt{1-\rho^2}}\left(\frac{r-\mu_2}{\sigma_2} - \rho\frac{r-\mu_1}{\sigma_1}\right)t$$

Note that W_t^{*1} and W_t^{*2} are independent. The price of the product call option is the discounted expected value (under Q) of the future payoff:

$$C(S_t^1 S_t^2, \tau) = \exp(-r\tau)\mathsf{E}^Q\left[(S_T^1 S_T^2 - K)^+ | \mathcal{F}_t\right]$$

As

$$(S_T^1 S_T^2 - K)^+ = \begin{cases} S_T^1 S_T^2 - K & \text{if } S_T^1 S_T^2 - K > 0 \\ 0 & \text{if } S_T^1 S_T^2 - K \le 0 \end{cases}$$

the inequality $S_T^1 S_T^2 > K$ has to be considered from a stochastic (Q) point of view.

The product $S_t^1 S_t^2$ is calculated:

$$\begin{aligned} S_t^1 S_t^2 &= S_0^1 S_0^2 \exp\left\{\left(\mu_1 - \frac{\sigma_1^2}{2}\right)t + \sigma_1 W_t^1 + \left(\mu_2 - \frac{\sigma_2^2}{2}\right)t \right. \\ &\quad \left. + \sigma_2(\rho W_t^1 + \sqrt{1-\rho^2}W_t^2)\right\} \\ &= S_0^1 S_0^2 \exp\left[\left(\mu_1 + \mu_2 - \frac{\sigma_1^2+\sigma_2^2}{2}\right)t + \sigma_1\left(W_t^{1*} + \theta_1 t\right) \right. \\ &\quad \left. + \sigma_2\left\{\rho(W_t^{1*} + \theta_1 t) + \sqrt{1-\rho^2}\left(W_t^{2*} + \theta_2 t\right)\right\}\right] \\ &= S_0^1 S_0^2 \exp\left\{\left(2r - \frac{\sigma_1^2+\sigma_2^2}{2}\right)t + (\sigma_1 + \sigma_2\rho)\, W_t^{1*} + \sigma_2\sqrt{1-\rho^2}W_t^{2*}\right\} \\ &= S_0^1 S_0^2 \exp\left\{\left(2r - \frac{\sigma_1^2+\sigma_2^2}{2}\right)t + \tilde{W}_t\right\} \end{aligned}$$

where $\tilde{W}_t = (\sigma_1 + \sigma_2\rho)\, W_t^{1*} + \sigma_2\sqrt{1-\rho^2}W_t^{2*} \sim \mathrm{N}(0, \tilde{\sigma}^2 t)$, and $\tilde{\sigma}^2 = \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2$.

It follows

$$S_T^1 S_T^2 = S_t^1 S_t^2 \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{W}_T - \tilde{W}_t\right\} \tag{9.8}$$

and therefore

$$S_T^1 S_T^2 > K$$

being equivalent to

$$Z > \frac{\log\left(\frac{K}{S_t^1 S_t^2}\right) - \left(2r - \frac{\sigma_1^2}{2} - \frac{\sigma_2^2}{2}\right)\tau}{\sqrt{\tau}\,\tilde{\sigma}} \stackrel{\text{def}}{=} -d_1 \tag{9.9}$$

for $Z \sim \mathrm{N}(0, 1)$ under Q.

Finally with (9.8) and (9.9):

$$\begin{aligned}
C(S_t^1 S_t^2, \tau) &= \exp(-r\tau)\mathsf{E}^Q\left[(S_T^1 S_T^2 - K)^+ | \mathcal{F}_t\right] \\
&= \exp(-r\tau)\mathsf{E}^Q\left[\left(S_t^1 S_t^2 \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{W}_T - \tilde{W}_t\right\} - K\right)^+ \Big| \mathcal{F}_t\right] \\
&= \exp(-r\tau)\mathsf{E}^Q\left[\left(S_t^1 S_t^2 \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\, Z\right\} - K\right)^+ \Big| \mathcal{F}_t\right] \\
&= \exp(-r\tau)\int_{-\infty}^{\infty}\left(S_t^1 S_t^2 \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\, x\right\} - K\right)^+ \varphi(x)\, dx \\
&= \exp(-r\tau)\int_{-d_1}^{\infty}\left(S_t^1 S_t^2 \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\, x\right\} - K\right) \varphi(x)\, dx \\
&= \exp(-r\tau)\frac{1}{\sqrt{2\pi}}\left[\int_{-d_1}^{\infty} S_t^1 S_t^2 \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\, x - \frac{x^2}{2}\right\} dx\right. \\
&\quad \left. - \int_{-d_1}^{\infty} K \exp\left(-\frac{x^2}{2}\right) dx\right] \\
&= \exp(-r\tau)\frac{1}{\sqrt{2\pi}}\left[S_t^1 S_t^2 \int_{-d_1}^{\infty} \exp\left\{\left(2r - \frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\, x - \frac{x^2}{2}\right\} dx\right. \\
&\quad \left. - K \int_{-\infty}^{d_1} \exp\left(-\frac{x^2}{2}\right) dx\right] \\
&= \exp(r\tau)\frac{1}{\sqrt{2\pi}} S_t^1 S_t^2 \int_{-d_1}^{\infty} \exp\left\{\left(-\frac{\sigma_1^2 + \sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\, x - \frac{x^2}{2}\right\} dx \\
&\quad - \exp(-r\tau) K \Phi(d_1)
\end{aligned} \tag{9.10}$$

Let $y = x - \tilde{\sigma}\sqrt{\tau}$.

$$\int_{-d_1}^{\infty} \exp\left\{\left(-\frac{\sigma_1^2+\sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\,x - \frac{x^2}{2}\right\} dx$$

$$= \int_{-d_1-\tilde{\sigma}\sqrt{\tau}}^{\infty} \exp\left\{\left(-\frac{\sigma_1^2+\sigma_2^2}{2}\right)\tau + \tilde{\sigma}\sqrt{\tau}\left(y + \tilde{\sigma}\sqrt{\tau}\right) - \frac{\left(y+\tilde{\sigma}\sqrt{\tau}\right)^2}{2}\right\} dy$$

$$= \int_{-d_2}^{\infty} \exp\left(\frac{\tilde{\sigma}^2 - \sigma_1^2 - \sigma_2^2}{2}\tau - \frac{y^2}{2}\right) dy$$

$$= \exp\left(\frac{\tilde{\sigma}^2 - \sigma_1^2 - \sigma_2^2}{2}\tau\right) \int_{-d_2}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy$$

$$= \exp\left(\sigma_1\sigma_2\rho\tau\right) \int_{-\infty}^{d_2} \exp\left(-\frac{y^2}{2}\right) dy$$

Thus (9.10) reads:

$$\exp(r\tau)S_t^1 S_t^2 \exp(\sigma_1\sigma_2\rho\tau)\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} \exp\left(-\frac{y^2}{2}\right) dy - \exp(-r\tau)K\Phi\left(d_1\right)$$
$$= \exp\{(r+\sigma_1\sigma_2\rho)\tau\}S_t^1 S_t^2 \Phi\left(d_2\right) - \exp(-r\tau)K\Phi\left(d_1\right)$$

The price of the European Product Call is found finally as:

$$C(S_t^1 S_t^2, \tau) = \exp\{(r+\sigma_1\sigma_2\rho)\tau\}\, S_t^1 S_t^2\, \Phi\left(d_2\right) - \exp(-r\tau)K\Phi\left(d_1\right) \tag{9.11}$$

$$d_1 = \frac{\log\left(\frac{S_t^1 S_t^2}{K}\right) + \left(2r - \frac{\sigma_1^2+\sigma_2^2}{2}\right)\tau}{\sqrt{\tau}\,\tilde{\sigma}}$$

$$d_2 = d_1 + \tilde{\sigma}\sqrt{\tau}$$

Exercise 9.8 (Product Call Option). *Consider two stock prices, one for Allianz* $S_t^1 = 60$ *and one for Munich Re* $S_t^2 = 100$ *with volatilities* $\sigma_1 = 42.49$ *and* $\sigma_2 = 31.4$ *and we assume that both stocks are correlated with coefficient* $\rho = 0.3$. *Calculate the price of the product call option if the interest rate is* 1 %, *strike price* $K = 6{,}000$ *and time to maturity is 1 year.*

According to the formulas given in Exercise 9.7:

$$\tilde{\sigma}^2 = 0.4249^2 + 2\cdot 0.3\cdot 0.4249\cdot 0.314 + 0.314^2$$
$$= 0.3592$$

$$d_1 = \frac{\log\left(\frac{60\cdot 100}{6000}\right) + \left(2\cdot 0.01 - \frac{0.4249^2+0.314^2}{2}\right)\cdot 1}{\sqrt{1}\,\sqrt{0.3582}}$$

$$
\begin{aligned}
&= -0.1995 \\
d_2 &= -0.1995 + \sqrt{0.3582}\sqrt{1} \\
&= 0.3998
\end{aligned}
$$

Thus the price of the European Product Call is:

$$
\begin{aligned}
C(60 \cdot 100, 1) = &\exp\{(0.01 + 0.4249 \cdot 0.314 \cdot 0.3) \cdot 1\} \, 60 * 100 \ \cdot \Phi\,(0.3998) \\
&- \exp(-0.01 \cdot 1) \cdot 6000 \cdot \Phi\,(-0.1995) \\
= &\ 1633.364
\end{aligned}
$$

SFSproductcall

Exercise 9.9 (Option pricing on an arithmetic Brownian motion). *Derive the pricing formula for a call option on S_t, where S_t follows an arithmetic Brownian motion. It is assumed that the riskless interest rate $r = 0$, the stock pays no dividends and the option is* at-the-money. *Calculate the distribution of S_t, the pricing formula on S_t and give the call option price.*

An arithmetic Brownian motion S_t with drift μ and volatility σ is described as:

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

where W_t is a standard Brownian motion.

Consider an American call option given by

$$C = \max\{S_t - K, 0\}$$

Under the risk-neutral measure Q the price of the call option is:

$$C(S, \tau) = \mathsf{E}^Q \{\exp(-r\tau) \max(S_\tau - K, 0)\}$$

with interest rate r. Now, we introduce

$$\tilde{W}_t = r - \mu + W_t.$$

Concerning the distribution of S_t with $\mu = r = 0$, we now have:

$$dS_t = r\ dt + \sigma\ d\tilde{W}_t = \sigma\ d\tilde{W}_t$$

Integrating yields:

$$S_t = S_0 + \sigma\ \tilde{W}_t$$

Hence,

$$S_t \sim \mathrm{N}(S_0, \sigma^2 t).$$

Calculating the call option price reads

$$\begin{aligned}
C(S,\tau) &= \mathsf{E}^{Q}[\max\{S_\tau - K, 0\}] \\
&= \mathsf{E}^{Q}[(S_\tau - K)\mathbf{1}(S_\tau > K)] \\
&= \int_K^\infty (S_\tau - K)\frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left\{-\frac{(S_\tau - S_0)^2}{2\sigma^2\tau}\right\} dS_\tau \\
&= \int_K^\infty S_\tau \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left\{-\frac{(S_\tau - S_0)^2}{2\sigma^2\tau}\right\} dS_\tau \\
&\quad - \int_K^\infty K \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left\{-\frac{(S_\tau - S_0)^2}{2\sigma^2\tau}\right\} dS_\tau
\end{aligned}$$

For the first term

$$= \int_K^\infty S_\tau \frac{1}{\sqrt{2\pi\sigma^2\tau}} \exp\left\{-\frac{(S_\tau - S_0)^2}{2\sigma^2\tau}\right\} dS_\tau$$

we define $y = \dfrac{S_\tau - S_0}{\sqrt{\sigma^2\tau}}$ to obtain:

$$\int_{\frac{K-S_0}{\sqrt{\sigma^2\tau}}}^\infty (S_0 + \sqrt{\sigma^2\tau}\, y)\sqrt{\frac{\sigma^2\tau}{2\pi\sigma^2\tau}} \exp\left(-\frac{y^2}{2}\right) dy$$

$$\begin{aligned}
&= \sqrt{\frac{\sigma^2\tau}{2\pi}} \int_{\frac{K-S_0}{\sqrt{\sigma^2\tau}}}^\infty y \exp\left(-\frac{y^2}{2}\right) dy + S_0 \int_{\frac{K-S_0}{\sqrt{\sigma^2\tau}}}^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\
&= \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} \left[-\exp\left(-\frac{y^2}{2}\right)\right]_{\frac{K-S_0}{\sqrt{\sigma^2\tau}}}^\infty + S_0 \Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right) \\
&= \sigma\sqrt{\tau}\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(K - S_0)^2}{2\sigma^2\tau}\right\} + S_0 \Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right) \\
&= \sigma\sqrt{\tau}\varphi\left(\frac{K - S_0}{\sigma\sqrt{\tau}}\right) + S_0 \Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right)
\end{aligned}$$

The second term involves the distribution of S_τ. Hence, we have

$$\begin{aligned}
K\ Q(S_\tau > K) &= K\ Q\left(\frac{S_\tau - S_0}{\sigma\sqrt{\tau}} > \frac{K - S_0}{\sigma\sqrt{\tau}}\right) \\
&= K\ \left\{1 - Q\left(\frac{S_\tau - S_0}{\sigma\sqrt{\tau}} \le \frac{K - S_0}{\sigma\sqrt{\tau}}\right)\right\}
\end{aligned}$$

$$= K \left\{1 - \Phi\left(\frac{K - S_0}{\sigma\sqrt{\tau}}\right)\right\}$$
$$= K\,\Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right)$$

Finally, the call option price reads

$$C(S,\tau) = \sigma\sqrt{\tau}\varphi\left(\frac{K - S_0}{\sigma\sqrt{\tau}}\right) + S_0\Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right) - K\Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right)$$
$$= \sigma\sqrt{\tau}\varphi\left(\frac{K - S_0}{\sigma\sqrt{\tau}}\right) + (S_0 - K)\Phi\left(\frac{S_0 - K}{\sigma\sqrt{\tau}}\right)$$

Exercise 9.10 (Power Call Option). *Let us consider a Power call option with payoff structure* $(S_T^\alpha - K)^+ = \max\{S_T^\alpha - K, 0\}$. *This is an example of a nonlinear payoff function, of which closed-form solutions are available. A higher payoff is possible compared to plain vanilla options, but also a higher premium compared to plain vanilla options is to be expected. Calculate the fair price for such a Power call option. Apply the Black-Scholes assumptions, i.e. suppose that* S_t *follows a geometric Brownian motion.*

The SDE of the underlying asset is:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{9.12}$$

The solution is

$$S_T = S_t \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)\tau + \sigma(W_T - W_t)\right\}$$

or

$$S_T = S_t \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)\tau + \sigma\sqrt{\tau}Z\right\},\ Z \sim \mathrm{N}(0,1)$$

Under the risk-neutral pricing measure measure Q, the underlying asset evolves according to

$$dS_t = rS_t dt + \sigma S_t d\tilde{W}_t \qquad \text{where } \tilde{W}_t \text{ is a } Q\text{–Brownian Motion.}$$

The existence of the Q–Brownian Motion $\tilde{W}_t$ is ensured by the Girsanov Theorem 22.4 in Franke et al. (2011).

The solution of (9.12) under Q is

$$S_T = S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau + \sigma(\tilde{W}_T - \tilde{W}_t)\right\}$$

or

$$S_T = S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)\tau + \sigma\sqrt{\tau}Z\right\}$$

Pricing formula for a contingent claim C with $\tau = T - t$:

$$C(S_t, \tau) = \exp(-r\tau)\mathsf{E}^Q[C(S_T, 0)|\mathcal{F}_t]$$

where $C(S_T, 0)$ is the payoff function of the contingent claim at maturity of the claim. Thus we need to evaluate

$$C(S_t, \tau) = \exp(-r\tau)\mathsf{E}^Q[(S_T^\alpha - K)^+|\mathcal{F}_t]$$

The conditional expectation $\mathsf{E}^Q[(S_T^\alpha - K)^+|\mathcal{F}_t]$ can be rewritten as

$$\mathsf{E}^Q[(S_T^\alpha - K)^+|\mathcal{F}_t] = \mathsf{E}^Q[S_T^\alpha \mathbf{1}(S_T^\alpha > K)|\mathcal{F}_t] - K\mathsf{E}^Q[\mathbf{1}(S_T^\alpha > K)|\mathcal{F}_t] = T_1 - T_2$$

For the dynamics of S_t^α, we apply Itô's formula with $S_t^\alpha = f(S_t, t)$ and (9.5):

$$\frac{\partial f(S_t, t)}{\partial t} = 0 \qquad \frac{\partial f(S_t, t)}{\partial s} = \alpha S_t^{\alpha-1} \qquad \frac{\partial^2 f(S_t, t)}{\partial s^2} = (\alpha - 1)\alpha S_t^{\alpha-2}$$

Thus

$$\begin{aligned} dS_t^\alpha &= \frac{1}{2}(\alpha - 1)\alpha S_t^{\alpha-2}\sigma^2 S_t^2 dt + \alpha S_t^{\alpha-1} dS_t \\ &= \frac{1}{2}(\alpha - 1)\alpha S_t^\alpha \sigma^2 dt + \alpha S_t^{\alpha-1}(\mu S_t dt + \sigma S_t dW_t) \\ &= \underbrace{\left\{\alpha\mu + \frac{1}{2}\sigma^2\alpha(\alpha - 1)\right\}}_{\tilde{\mu}} S_t^\alpha dt + \underbrace{\alpha\sigma}_{\tilde{\sigma}} S_t^\alpha dW_t \end{aligned}$$

The solution is:

$$\begin{aligned} S_T^\alpha &= S_t^\alpha \exp\left\{\left(\tilde{\mu} - \frac{1}{2}\tilde{\sigma}^2\right)\tau + \tilde{\sigma}(W_T - W_t)\right\} \\ &= S_t^\alpha \exp\left[\left\{\alpha\mu + \frac{1}{2}\sigma^2\alpha(\alpha - 1) - \frac{1}{2}\alpha^2\sigma^2\right\}\tau + \alpha\sigma(W_T - W_t)\right] \\ &= S_t^\alpha \exp\left\{\alpha\left(\mu - \frac{1}{2}\sigma^2\right)\tau + \alpha\sigma(W_T - W_t)\right\} \end{aligned}$$

Under Q, the term in T_1 equals:

$$S_T^\alpha = S_t^\alpha \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \alpha\sigma(\tilde{W}_T - \tilde{W}_t)\right\}$$

or

$$S_T^\alpha = S_t^\alpha \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \alpha\sigma\sqrt{\tau}Z\right\}$$

For the evaluation of the second term T_2 we find an expression for:

$$\mathsf{E}^Q[\mathbf{1}(S_T^\alpha > K)|\mathcal{F}_t]$$

Note that $S_T^\alpha > K$ when $\log(S_T^\alpha) > \log(K)$ or:

$$\log(S_t^\alpha) + \alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \alpha\sigma\sqrt{\tau}Z > \log(K)$$

$$\log(S_t) + \left(r - \frac{1}{2}\sigma^2\right)\tau + \sigma\sqrt{\tau}Z > \frac{1}{\alpha}\log(K)$$

$$\frac{\frac{1}{\alpha}\log(K) - \log(S_t) - \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} < Z$$

$$\frac{\log(K^{\frac{1}{\alpha}}/S_t) - \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} \stackrel{\text{def}}{=} -z^* < Z$$

Hence

$$\mathsf{E}^Q[\mathbf{1}_{\{S_T^\alpha > K\}}|\mathcal{F}_t] = Q\big(Z > -z^*\big) = \mathrm{P}(Z > -z^*) = \int_{-z^*}^{\infty} \varphi(z)dz$$
$$= \int_{-\infty}^{z^*} \varphi(z)dz = \Phi(z^*)$$

where $\varphi(z)$ and $\Phi(z)$ are the density and distribution function of the standard normal variable Z, respectively.

Thus:

$$\mathsf{E}^Q[\mathbf{1}(S_T^\alpha > K)|\mathcal{F}_t] = \Phi(z^*)$$

For T_1 we obtain:

$$\mathsf{E}^Q\left[S_T^\alpha \mathbf{1}(S_T^\alpha > K)|\mathcal{F}_t\right] = \frac{1}{\sqrt{2\pi}} \int_{-z^*}^{\infty} \overbrace{S_t^\alpha \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \alpha\sigma\sqrt{\tau}z\right\}}^{S_T^\alpha}$$
$$\cdot \exp\left(-\frac{1}{2}z^2\right)dz$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \int_{-z^*}^{\infty} S_t^{\alpha} \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \alpha\sigma\sqrt{\tau}z - \frac{1}{2}z^2\right\} dz \\
&= \frac{1}{\sqrt{2\pi}} \underbrace{S_t^{\alpha} \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau\right\}}_{:=X_t} \int_{-z^*}^{\infty} \exp\left(\alpha\sigma\sqrt{\tau}z - \frac{1}{2}z^2\right) dz
\end{aligned} \tag{9.13}
$$

Manipulation of the integrand yields:

$$
\begin{aligned}
-\frac{(z-\alpha\sigma\sqrt{\tau})^2}{2} &= -\frac{z^2 - 2\alpha\sigma\sqrt{\tau}z + \alpha^2\sigma^2\tau}{2} \\
&= -\frac{1}{2}z^2 + \alpha\sigma\sqrt{\tau}z - \frac{1}{2}\alpha^2\sigma^2\tau
\end{aligned}
$$

Thus:

$$
-\frac{(z-\alpha\sigma\sqrt{\tau})^2}{2} + \frac{1}{2}\alpha^2\sigma^2\tau = -\frac{1}{2}z^2 + \alpha\sigma\sqrt{\tau}z
$$

Insert this into (9.13):

$$
\begin{aligned}
T_1 &= \frac{1}{\sqrt{2\pi}} X_t \int_{-z^*}^{\infty} \exp\left(-\frac{1}{2}z^2 + \alpha\sigma\sqrt{\tau}z\right) dz \\
&= \frac{1}{\sqrt{2\pi}} X_t \int_{-z^*}^{\infty} \exp\left\{-\frac{(z-\alpha\sigma\sqrt{\tau})^2}{2} + \frac{1}{2}\alpha^2\sigma^2\tau\right\} dz \\
&= \frac{1}{\sqrt{2\pi}} X_t \exp\left(\frac{1}{2}\alpha^2\sigma^2\tau\right) \int_{-z^*}^{\infty} \exp\left\{-\frac{1}{2}(z-\alpha\sigma\sqrt{\tau})^2\right\} dz
\end{aligned} \tag{9.14}
$$

Now, define $y = z - \alpha\sigma\sqrt{\tau}$ and observe $dy = dz$ and the lower bound of integral changes from $-z^*$ to $(-z^* - \alpha\sigma\sqrt{\tau})$

Thus (9.14) reads:

$$
\begin{aligned}
T_1 &= X_t \exp\left(\frac{1}{2}\alpha^2\sigma^2\tau\right) \int_{-z^*-\alpha\sigma\sqrt{\tau}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) dy \\
&= X_t \exp\left(\frac{1}{2}\alpha^2\sigma^2\tau\right) \int_{-z^*-\alpha\sigma\sqrt{\tau}}^{\infty} \varphi(y)dy \\
&= X_t \exp\left(\frac{1}{2}\alpha^2\sigma^2\tau\right) \int_{-\infty}^{z^*+\alpha\sigma\sqrt{\tau}} \varphi(y)dy \qquad \text{(symmetry)} \\
&= X_t \exp\left(\frac{1}{2}\alpha^2\sigma^2\tau\right) \Phi\left(z^* + \alpha\sigma\sqrt{\tau}\right)
\end{aligned}
$$

Recall now $X_t = S_t^\alpha \exp\{\alpha \left(r - \frac{1}{2}\sigma^2\right) \tau\}$, thus:

$$\begin{aligned}
&X_t \cdot \exp\left(\frac{1}{2}\alpha^2\sigma^2\tau\right) \Phi\left(z^* + \alpha\sigma\sqrt{\tau}\right)\\
&= S_t^\alpha \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \frac{1}{2}\alpha^2\sigma^2\tau\right\} \Phi\left(z^* + \alpha\sigma\sqrt{\tau}\right)
\end{aligned}$$

Thus:

$$\begin{aligned}
\mathsf{E}^Q[S_T^\alpha \mathbf{1}(S_T^\alpha > K)|\mathcal{F}_t] = S_t^\alpha \exp\left\{\alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \frac{1}{2}\alpha^2\sigma^2\tau\right\}\\
\cdot \Phi\left(z^* + \alpha\sigma\sqrt{\tau}\right)
\end{aligned}$$

Discounting the summands Risk–neutral pricing equation:

$$\begin{aligned}
C_t &= \exp(-r\tau)\,\mathsf{E}^Q[S_T^\alpha - K|\mathcal{F}_t]\\
&= \exp(-r\tau)\,T_1 - K\exp(-r\tau)\,T_2
\end{aligned}$$

Discounting the first summand over the remaining lifetime of the option yields:

$$\begin{aligned}
&S_t^\alpha \exp\left\{-r\tau + \alpha\left(r - \frac{1}{2}\sigma^2\right)\tau + \frac{1}{2}\alpha^2\sigma^2\tau\right\} \Phi(\cdot)\\
&= S_t^\alpha \exp\left\{(\alpha - 1)\left(r + \frac{1}{2}\alpha\sigma^2\right)\tau\right\} \Phi(z^* + \alpha\sigma\sqrt{\tau})
\end{aligned}$$

Discounting the second summand yields: $\exp(-r\tau)\,\Phi(z^*)$.

Putting pieces together yields the pricing equation for a European Power Call option:

$$C_t = S_t^\alpha \exp\left\{(\alpha - 1)\left(r + \frac{1}{2}\alpha\sigma^2\right)\tau\right\} \Phi\left(z^* + \alpha\sigma\sqrt{\tau}\right) - K\exp(-r\tau)\,\Phi(z^*) \tag{9.15}$$

$$\text{with} \quad z^* + \alpha\sigma\sqrt{\tau} = \frac{\log(S_t/K^{\frac{1}{\alpha}}) + \left\{r + (\alpha - \frac{1}{2})\sigma^2\right\}\tau}{\sigma\sqrt{\tau}}$$

$$z^* = \frac{\log(S_t/K^{\frac{1}{\alpha}}) + \left(r - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}$$

Table 9.1 Comparison of BS plain vanilla call and BS power call. `SFScompBS`

	$S_t = 10$ (ATM vanilla)	$S_t = 15$ (ITM vanilla)
Vanilla call	1.21	5.37
Power call	101.10	239.61

Exercise 9.11 (Plain Vanilla Call Option and BS Power Call Option). *Let us compare (9.15) with a plain vanilla BS price. Choose $K = 10$, $r = 0.02$, $\sigma = 0.22$, $\tau = 1.5$, and $\alpha = 2$. Calculate the price for a BS plain vanilla call option and a BS power call option under $S_t = 10$ and $S_t = 15$.*

Plugging the given parameters into the above stated definition of z^* yields

$$z^* = \frac{\log(S_t/10^{\frac{1}{2}}) + \left(0.02 - 0.5 \cdot 0.22^2\right) 1.5}{0.22\sqrt{1.5}}$$

Hence, one obtains $z^*(S_t = 10) = 4.25726$ and $z^*(S_t = 15) = 5.762082$. Thus, the price of the BS power call option under $S_t = 10$ is

$$\begin{aligned} C_t &= 10^2 \cdot \exp\left\{(2-1)\left(0.02 + 0.5 \cdot 2 \cdot 0.22^2\right) \cdot 1.5\right\} \cdot \Phi(4.25726 + 2 \cdot 0.22\sqrt{1.5}) \\ &\quad - 10 \cdot \exp(-0.02 \cdot 1.5)\Phi(4.25726) \\ &= 101.1004 \end{aligned}$$

Applying the same procedure for $S_t = 15$ yields $C_t = 239.6064$.

The BS formula for a plain vanilla European call option reads:

$$C(S_t, \tau) = S_t \Phi(d_1) - K \exp(r\tau)\Phi(d_2),$$

where we use the abbreviations

$$\begin{aligned} d_1 &= y + \sigma\sqrt{\tau}, \\ d_2 &= y \\ y &= \frac{\log S_t/K + \left(r - \sigma^2/2\right)\tau}{\sigma\sqrt{\tau}}. \end{aligned}$$

Plugging in the given parameters yields $C_t = 1.210116$ for $S_t = 10$ and 5.367782 for $S_t = 15$ (Table 9.1).

Chapter 10
Models for the Interest Rate and Interest Rate Derivatives

利率和利率衍生品模型

天有不测風雲，人有旦夕禍福。
Human fortunes are as unpredictable as the weather.

Pricing interest rate derivatives fundamentally depends on the underlying term structure. The often made assumptions of constant risk free interest rate and its independence of equity prices will not be reasonable when considering interest rate derivatives. Just as the dynamics of a stock price are modeled via a stochastic process, the term structure of interest rates is modeled stochastically. As interest rate derivatives have become increasingly popular, especially among institutional investors, the standard models for the term structure have become a core part of financial engineering. It is therefore important to practice the basic tools of pricing interest rate derivatives. For interest rate dynamics, there are one-factor and two-factor short rate models, the Heath Jarrow Morton framework and the LIBOR Market Model.

Exercise 10.1 (Forward Rate Agreements and Receiver Interest Rate Swap). *Consider the setup in Table 10.1 with the face value of the considered bonds as 1 EUR.*

(a) Calculate the value of the forward rate agreements.
(b) Calculate the value of a receiver interest rate swap.
(c) Determine the swap rate.

(a) A forward rate agreement $FRA_{R_K,S}\{r(t), t, T\}$ is an agreement at time t that a certain interest rate R_K will apply to a principal amount for a certain period of time $\tau(T, S)$, in exchange for an interest rate payment at the future interest rates $R(T, S)$, with $t < T < S$.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_10,

Table 10.1 Dataset

Maturity(years)	0.5	1	1.5	2
Bond value	0.97	0.94	0.91	0.87
Strike rate (%)	7.50	7.50	7.50	7.50

The value of a forward rate agreement is determined by:

$$\begin{aligned} \text{FRA}_{R_K,S}\{r(t),t,T\} &= \frac{\tau(T,S)\{R_K - R(T,S)\}}{1 + R(t,S)\tau(t,S)} \\ &= V(t,S)\tau(T,S)R_K + V(t,S) - V(t,T) \qquad (10.1) \end{aligned}$$

where t is the current time, the time when FRAs come into place is T, and the maturity of the FRAs is S. Here R_K stands for the strike interest rate.

The term structure of interest rates is therefore not needed. $\tau(T,S) = 0.5$ for all FRAs. Plug in (10.1), we now calculate:

$$\begin{aligned} \text{FRA}_{0.075,0.5}\{r(t),0,0.0\} &= 0.97 \cdot 0.5 \cdot 0.075 + 0.97 - 1.00 = 0.0064 \\ \text{FRA}_{0.075,1.0}\{r(t),0,0.5\} &= 0.94 \cdot 0.5 \cdot 0.075 + 0.94 - 0.97 = 0.0053 \\ \text{FRA}_{0.075,1.5}\{r(t),0,0.5\} &= 0.91 \cdot 0.5 \cdot 0.075 + 0.91 - 0.94 = 0.0041 \\ \text{FRA}_{0.075,2.0}\{r(t),0,0.5\} &= 0.87 \cdot 0.5 \cdot 0.075 + 0.87 - 0.91 = -0.0074 \end{aligned}$$

These results are given in Table 10.2.

(b) An Interest Rate Swap $\text{IRS}_{R_K,T}\{r(t),t\}$ is an agreement to exchange payments of a fixed rate R_K against a variable rate $R(t,t_i)$ over a period $\tau(t,T)$ at certain time points t_i, with $t \le t_i \le T$. When we consider a receiver interest rate swap, we receive the fixed interest rate in exchange for paying the floating rate.

For the valuation of the receiver interest rate swap, we can apply two different methods. First, we can value the fixed leg and floating leg of the swap separately. This would correspond to thinking of an IRS as an agreement to exchange a coupon-bearing bond for a floating rate note. For the fixed leg we set the coupon payments equal to:

$$c_i = \tau_i R_K,$$

where τ_i is the time to maturity of bond i. This gives us:

$$\begin{aligned} \text{FixedLeg}_{R_K}\{r(t),t\} &= \sum_{i=0}^{n-1}\{1 + R(t,t_{i+1})\tau_i\}^{-1} c_i + V(t,T) \\ &= \sum_{i=0}^{n-1} V(t,t_{i+1}) R_K \tau_i + V(t,T) \end{aligned}$$

Table 10.2 Forward rate agreements

Maturity (years)	0.5	1	1.5	2
FRA	0.0064	0.0053	0.0041	−0.0074

$$
\begin{aligned}
&= 0.97 \cdot 0.075 \cdot 0.5 + 0.94 \cdot 0.075 \cdot 0.5 + 0.91 \cdot 0.075 \cdot 0.5 \\
&\quad +0.87 \cdot 0.075 \cdot 0.5 + 0.87 \\
&= 1.008375
\end{aligned}
$$

The value of the floating leg will, by definition, always equal to 1 EUR. Thus the value of the receiver interest rate swap equals to:

$$
\begin{aligned}
\text{RIRS}_{0.075,2}\{r(t),0\} &= \text{FixedLeg}\{r(t),0\} - \text{FloatingLeg}\{r(t),0\} \\
&= 1.008375 - 1 \\
&= 0.008375
\end{aligned}
$$

Alternatively, we can value the swap by adding the values of the seperate FRAs from Table 10.2:

$$
\begin{aligned}
\text{RIRS}_{0.075,2}\{r(t),0\} &= \sum_{i=0}^{n-1} \text{FRA}_{0.075,t_{i+1}}\{r(t),0,t_i\} \\
&= 0.0064 + 0.0053 + 0.0041 - 0.0074 \\
&= 0.0084
\end{aligned}
$$

(c) The swap rate is:

$$
\begin{aligned}
R_S(0,2) &= \{1 - V(0,2)\} / \left\{ \sum_{i=0}^{n-1} V(0,t_{i+1})\tau_i \right\} \\
&= (1-0.87)/(0.97 \cdot 0.5 + 0.94 \cdot 0.5 + 0.91 \cdot 0.5 + 0.87 \cdot 0.5) \\
&= 0.07
\end{aligned}
$$

We can also calculate the swap rate by setting the value of the receiver interest rate swap equal to zero:

$$
\begin{aligned}
\text{RIRS}_{R_S,2}\{r(t),0\} &= 0 \\
0.97 \cdot R_S \cdot 0.5 + 0.94 \cdot R_S \cdot 0.5 & \\
+0.91 \cdot R_S \cdot 0.5 + 0.87 \cdot R_S \cdot 0.5 + 0.87 - 1 &= 0 \\
1.845 \cdot R_S - 0.13 &= 0 \\
R_S &= 0.07
\end{aligned}
$$

Exercise 10.2 (Heath Jarrow Morton Framework). *Consider the one factor Heath Jarrow Morton framework:*

$$df(t,T) = \alpha(t,T)dt + \beta(t,T)dW_t \tag{10.2}$$

where

$$\alpha(t,T) = \sigma(t,T)\left\{\frac{\partial\sigma(t,T)}{\partial T}\right\}$$

and

$$\beta(t,T) = \left\{\frac{\partial\sigma(t,T)}{\partial T}\right\}.$$

(a) Show that if volatility is constant, this model is reduced to the Ho-Lee model for the short rate dynamic process.
(b) Show that if volatility $\sigma(t,T) = \sigma\exp\{-a(T-t)\}$, where a and σ are constants, this model is reduced to the Hull-White model for the short rate dynamic process.

(a) The Ho-Lee model is:

$$dr(t) = \delta(t)dt + \sigma dW_t.$$

If $d\sigma(t,T)/dT = \sigma$, then:

$$\sigma(t,T) = \sigma \times (T-t).$$

Thus, the process for the instantaneous forward rate becomes:

$$\begin{aligned} f(t,T) &= f(0,T) + \sigma^2 \int_0^t (T-u)\,du + \sigma W_t \\ &= f(0,T) + \sigma^2(tT - \frac{t^2}{2}) + \sigma W_t. \end{aligned}$$

Since $r(t) = f(t,t)$:

$$r(t) = f(0,t) + \frac{\sigma^2 t^2}{2} + \sigma W_t$$

which gives the SDE:

$$dr(t) = \{f(0,t) + \sigma^2 t\}dt + \sigma dW_t.$$

So this model is the Ho-Lee model with $\delta(t) = f(0,t) + \sigma^2 t$ as a fixed function of time.

(b) The Hull-White model is:

$$dr(t) = \{\delta(t) - ar(t)\}dt + \sigma dW_t. \tag{10.3}$$

Here, we consider:

$$\frac{d\sigma(t,T)}{dT} = \sigma \exp\{-a(T-t)\}. \tag{10.4}$$

Therefore $\sigma(t,T)$ is equal to:

$$\sigma(t,T) = a^{-1}\sigma[1 - \exp\{-a(T-t)\}]. \tag{10.5}$$

Substituting (10.4) and (10.5) into (10.2):

$$df(t,T) = \frac{\sigma^2}{a}[1 - \exp\{-a(T-t)\}]\exp\{-a(T-t)\}dt + \sigma\exp\{-a(T-t)\}dW_t.$$

We get the instantaneous forward rate:

$$\begin{aligned} f(t,T) = {} & f(0,T) + \frac{1}{2}\sigma^2\{A(0,T)^2 - A(t,T)^2\} \\ & + \exp\{-a(T-t)\}\sigma\int_0^t \exp\{-a(t-s)\}\,dW_s \end{aligned}$$

where

$$A(t,T) = a^{-1}[1 - \exp\{-a(T-t)\}].$$

Again, let $r(t) = f(t,t)$, then the short rate is given by:

$$r(t) = f(0,t) + \frac{\sigma^2}{2a^2}\{1 - \exp(-at)\}^2 + \sigma\int_0^t \exp\{-a(t-s)\}\,dW_s.$$

The dynamic process for $r(t)$ is:

$$dr(t) = \left[\frac{\partial f}{\partial T}(0,t) + af(0,t) + \frac{\sigma^2}{2a}\{1 - \exp(-2at)\} - ar(t)\right] + \sigma dW_t.$$

This is equal to the Hull-White model in (10.3), with:

$$\delta(t) = \frac{\partial f}{\partial T}(0,t) + af(0,t) + \frac{\sigma^2}{2a}\{1 - \exp(-2at)\}$$

as a fixed function of time.

Exercise 10.3 (Hull-White Model). *Consider the Hull-White model*

$$\begin{aligned} dr(t) &= \mu_r dt + \sigma dW_t \\ \mu_r &= \delta(t) - ar \end{aligned} \tag{10.6}$$

Derive the price of zero-coupon bond at time t with a nominal face value of 1 *EUR under risk-neutral measures.*

Assume the bond value $V(r,t) = \exp\{A(t) - rB(t)\}$ and apply Itô's Lemma:

$$dV(r,t) = \frac{\partial V(r,t)}{\partial t}dt + \frac{1}{2}\sigma^2\frac{\partial^2 V(r,t)}{\partial r^2}dt + \frac{\partial V(r,t)}{\partial r}dr(t).$$

Plugging in the dynamic $r(t)$ to $V(r,t)$,

$$\begin{aligned} dV(r,t) = &\left\{\frac{\partial V(r,t)}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V(r,t)}{\partial r^2} + \mu_r\frac{\partial V(r,t)}{\partial r}\right\}dt \\ &+\sigma\frac{\partial V(r,t)}{\partial r}dW_t. \end{aligned} \tag{10.7}$$

Under risk-neutral measure, the market price of risk equals to zero, so the dynamic of the bond can be written as:

$$dV(r,t) = r(t)V(r,t)dt + \sigma_B V(r,t)dW_t, \tag{10.8}$$

We have from (10.7) and (10.8)

$$r(t)V(r,t) = \frac{\partial V(r,t)}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V(r,t)}{\partial r^2} + \mu_r\frac{\partial V(r,t)}{\partial r} \tag{10.9}$$

To solve this equation, according to

$$\begin{aligned} \frac{\partial V(r,t)}{\partial t} &= \{A'(t) - rB'(t)\}V(t) \\ \frac{\partial V(r,t)}{\partial r} &= -B(t)V(t) \\ \frac{\partial^2 V(r,t)}{\partial r^2} &= B^2(t)V(t) \end{aligned}$$

(10.9) becomes:

$$[A'(t) + \{aB(t) - B'(t) - 1\}r(t) - \{\delta(t) - \sigma^2 B(t)/2\}B(t)]V(t) = 0$$

then we get the solution:

$$B'(t) = aB(t) - 1$$
$$A'(t) = \{\delta(t) - \frac{1}{2}\sigma^2 B(t)\}B(t)$$

With the boundary condition $A(T) = B(T) = 0$ (since $V(r, T) = 1$), we can calculate A(t) and B(t) as:

$$\int_t^T A'(t) = \int_t^T \{\delta(t) - \frac{1}{2}\sigma^2 B(t)\}B(t)$$
$$A(t) = \int_t^T [\{-\delta(s) + \sigma^2 B(s)/2\}B(s)]\, ds$$
$$B'(t) - aB(t) = -1$$
$$\exp(-\int_0^s a\,du)B'(t) - a\exp(-\int_0^s a\,du)B(t) = -\exp(-\int_0^s a\,du)$$
$$d\{\exp(-\int_0^s a\,du)B(t)\}/\,ds = -\exp(-\int_0^s a\,du)$$
$$\int_t^T d\{\exp(-\int_0^s a\,du)B(t)\}/\,ds = \int_t^T \exp(-\int_0^s a\,du)\,ds$$
$$B(t) = \int_t^T \exp(-\int_t^s a\,du)\,ds$$

Moreover, there is an explicit solution

$$V(t) = \exp\{A(t) - rB(t)\}$$

with

$$A(t) = \int_t^T [\{-\delta(s) + \sigma^2 B(s)/2\}B(s)]\, ds$$
$$B(t) = \int_t^T \exp(-\int_t^s a\,du)\,ds$$

Exercise 10.4 (Pure-discount Bond in Vasicek Model). *Consider the Vasicek model*

$$dr(t) = a(b - r)dt + \sigma dW_t$$

where a, b, σ are known, W_t is a Wiener process. Derive the price of the pure-discount bond under real-world measure.

Assume the bond value $V(r,t) = \exp\{A(t) - rB(t)\}$. Similar to Exercise 10.3 we get:

$$dV(r,t) = \left\{\frac{\partial V(r,t)}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V(r,t)}{\partial r^2} + \mu_r \frac{\partial V(r,t)}{\partial r}\right\} dt + \sigma \frac{\partial V(r,t)}{\partial r} dW_t$$

with $\mu_r = a(b - r)$.
Under the real-world measure, the bond dynamic is

$$\begin{aligned} dV(r,t) &= \mu_B V(r,t)dt + \sigma_B V(r,t)dW_t \\ \mu_B V(r,t) &= \frac{\partial V(r,t)}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V(r,t)}{\partial r^2} + \mu_r \frac{\partial V(r,t)}{\partial r} \\ \sigma_B V(r,t) &= \sigma \frac{\partial V(r,t)}{\partial r} \end{aligned} \tag{10.10}$$

with $\mu_B \neq r(t)$. We use the market price of risk $\lambda(r,t)$ to represent μ_B. Different from Exercise 10.3, we define $w \stackrel{\text{def}}{=} \sigma$ here, then we have:

$$\begin{aligned} \lambda(r,t) &= \frac{\mu_B - r(t)}{\sigma_B} \\ \mu_B &= r(t) + \lambda_t w \frac{\partial V(r,t)}{\partial r} \end{aligned}$$

Under the real-world measure, (10.10) becomes:

$$\frac{\partial V(r,t)}{\partial t} + \frac{1}{2}w^2 \frac{\partial^2 V(r,t)}{\partial r^2} + (\mu_r - \lambda_t w)\frac{\partial V(r,t)}{\partial r} - r(t)V(r,t) = 0.$$

$$\begin{aligned} 0 &= \{A'(t) - B'(t)r(t)\}V(t) + \frac{1}{2}w^2 V(t)B^2(t) \\ &\quad - \{a(b-r) - \lambda_t w\}B(t)V(t) - r(t)V(t) \\ A'(t) &= (ab - \lambda_t w)B(t) - w^2 B^2(t)/2 \\ B'(t) &= aB(t) - 1 \end{aligned}$$

Assume $\lambda_t = \lambda$ and the boundary condition $A(T) = B(T) = 0$, there is an explicit solution:

$$V(r,t) = \exp\{A(t) - rB(t)\}$$

with

$$A(t) = (b - \lambda w/a - w^2/a^2)\{B(t) - T + t\} - w^2 B^2(t)/4a$$
$$B(t) = [1 - \exp\{-a(T - t)\}]/a$$

Exercise 10.5 (Vasicek Model). *Use the Vasicek model in Exercise 10.4,*

(a) Calculate $\mathsf{E}[r_t|\mathcal{F}_s]$ *and* $\mathsf{Var}[r_t|\mathcal{F}_s]$ *where* $s < t$ *and* $\mathcal{F}_s$ *denotes the past information set.*
(b) The yield to maturity is defined as:

$$Y_T(t) = -\log P_T(t)/\tau$$

where $\tau = T - t$, $P_T(t) = V(r,t)$. *Calculate the* $Y_{lim} = \lim_{\tau\to\infty} Y(\tau)$ *and what does it imply?*

(a)

$$dr(t) = ab\,dt - ar\,dt + \sigma\,dW_t$$
$$dr(t) + ar\,dt = ab\,dt + \sigma\,dW_t$$
$$d\exp(at)r(t) = \exp(at)(ab\,dt + \sigma\,dW_t)$$
$$\int_s^t d\exp(av)r(v)\,dv = ab\int_s^t \exp(av)\,dv + \sigma\int_s^t \exp(av)\,dW_v$$
$$\exp(at)r(t) - \exp(as)r(s) = b\exp(at) - b\exp as + \sigma\int_s^t \exp(av)dW_v$$

$$r(t) = \exp\{-a(t-s)\}r(s) + b[1-\exp\{-a(t-s)\}] + \sigma\exp(-at)\int_s^t \exp(av)dW_v$$

Since $\mathsf{E}[dW_t] = 0$, we have:

$$\mathsf{E}[r_t|\mathcal{F}_s] = b[1 - \exp\{-a(t-s)\}] + r(s)\exp\{-a(t-s)\}$$

According to Itô isometry: $\mathsf{E}\{\int_0^t f(s)dW_s\}^2 = \int_0^t \mathsf{E}\, f^2(s)ds$

$$\mathsf{Var}[r_t|\mathcal{F}_s] = \mathsf{E}[\sigma\exp(-at)\int_s^t \exp(av)dW_v]^2$$
$$= \sigma^2\exp(-2at)\int_s^t \exp(2av)dv$$
$$= \frac{\sigma^2}{2a}[1 - \exp\{-2a(t-s)\}]$$

(b)

$$
\begin{aligned}
Y_T(t) &= -\log P_T(t)/\tau = -\log V(r,t)/\tau \\
&= -\log \exp\{A(t) - rB(t)\}/\tau \\
Y_{lim} &= \lim_{\tau\to\infty} -\log V(r,t)/\tau = \lim_{\tau\to\infty} -\{A(t) - rB(t)\}/\tau \\
&= b - w\lambda/a - w^2/a^2
\end{aligned}
$$

When $\tau \to \infty$, the yield to maturity converges to a constant Y_{lim}. The bond value function can be rewritten as:

$$V(r,t) = \exp\{A(t) - rB(t)\}$$

with

$$
\begin{aligned}
A(t) &= Y_{lim}\{B(t) - \tau\} - w^2 B^2(t)/4a \\
B(t) &= [1 - \exp\{-a(T-t)\}]/a
\end{aligned}
$$

Part II
Statistical Model of Financial Time Series

Chapter 11
Financial Time Series Models

金融時系列モデル
急がば回れ
More haste, less speed.

This chapter deals with financial time series analysis. The statistical properties of asset and return time series are influenced by the media (daily news on the radio, television and newspapers) that inform us about the latest changes in stock prices, interest rates and exchange rates. This information is also available to traders who deal with immanent risk in security prices. It is therefore interesting to understand the behaviour of asset prices. Economic models on the pricing of securities are mostly based on theoretical concepts which involve the formation of expectations, utility functions and risk preferences. Here we concentrate on the empirical facts. Firstly, given a data set we aim to specify an appropriate model reflecting the main characteristics of the empirically observable stock price process and we wish to know whether the assumptions underlying the model are fulfilled in reality or whether the model has to be modified. A new model on the stock price process could possibly effect the function of the markets. To this end we apply statistical tools to empirical data and start with considering the concepts of univariate analysis before moving on to multivariate time series.

Exercise 11.1 (Stationarity and Autocorrelation). *Let X be a random variable with $\mathsf{E}(X^2) < \infty$ and define a stochastic process*

$$X_t \stackrel{\text{def}}{=} (-1)^t X, \ t = 1, 2, \ldots \tag{11.1}$$

(a) What do the paths of this process look like?
(b) Find a necessary and sufficient condition for X such that the process $\{X_t\}$ is strictly stationary.

DOI 10.1007/978-3-642-33929-5_11,

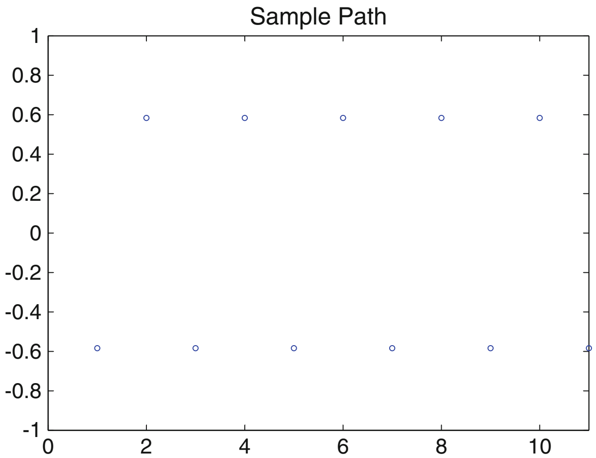


Fig. 11.1 Sample path for the case $X(\omega) = 0.5836$. SFSsamplepath

(c) *Find a necessary and sufficient condition for* X *such that* $\{X_t\}$ *is covariance (weakly) stationary.*
(d) *Let* X *be such that* $\{X_t\}$ *is covariance (weakly) stationary. Calculate the autocorrelation* ρ_τ*.*

(a) If for example $X(\omega) = 0.5836$, then the corresponding sample path is given in the Fig. 11.1.
(b) According to the definition, the stochastic process X_t is strictly stationary if for any $t_1, \dots, t_n$ and for all $n, s \in \mathbb{Z}$ it holds that

$$\begin{aligned} \mathrm{P}(X_{t_1} \le x_1, X_{t_2} \le x_2, \dots, X_{t_n} \le x_n) = \\ \mathrm{P}(X_{t_1+s} \le x_1, X_{t_2+s} \le x_2, \dots, X_{t_n+s} \le x_n). \end{aligned}$$

In our special case of the process $\{X_t\}$ defined by (11.1), the definition of strict stationarity reduces to

$$\mathrm{P}(X_1 \le a, X_2 \le b) = \mathrm{P}(X_2 \le a, X_3 \le b).$$

We check that this condition is fulfilled if, and only if, the distribution of X is symmetric, i.e. $\mathrm{P}(X \le x) = \mathrm{P}(-X \le x)$ for all x.

If the distribution of X is symmetric, it holds:

$$\begin{aligned} \mathrm{P}(X_1 \le a, X_2 \le b) &= \mathrm{P}(-X \le a, X \le b) \\ &= \mathrm{P}(-b \le -X \le a) \\ &= \mathrm{P}(-X \le a) - \mathrm{P}(-X < -b). \end{aligned}$$

Because of the symmetry of the distribution of X, we obtain:

$$\begin{aligned} P(-X \le a) - P(-X < -b) &= P(X \le a) - P(X < -b) \\ &= P(-b \le X \le a) \\ &= P(X \le a, -X \le b) \\ &= P(X_2 \le a, X_3 \le b) \end{aligned}$$

and thus, the process $\{X_t\}$ is strictly stationary. On the other hand, if we assume, that the process is strictly stationary, it holds:

$$P(X_1 \le a, X_2 \le b) = P(X_2 \le a, X_3 \le b).$$

Rewriting the last equation for our special case:

$$P(-X \le a, X \le b) = P(X \le a, -X \le b),$$

the symmetry of the distribution of X is obvious.

(c) The process $\{X_t\}$ is stationary if, and only if, $\mathsf{E}(X) = 0$.
If $\{X_t\}$ is stationary, it must hold:

$$\mathsf{E}(-X) = \mathsf{E}(X_1) = \mathsf{E}(X_2) = \mathsf{E}(X).$$

From $\mathsf{E}(-X) = \mathsf{E}(X)$ follows directly $\mathsf{E}(X) = 0$. If $\mathsf{E}(X) = 0$ then we obtain:

$$\begin{aligned} \mathsf{E}(X_t) &= \mathsf{E}\left\{(-1)^t X\right\} = (-1)^t\, \mathsf{E}(X) = 0 \\ \mathsf{Cov}\{X_t, X_{t+\tau}\} &= \mathsf{E}\left\{(-1)^t X (-1)^{t+\tau} X\right\} = (-1)^{2t+\tau}\, \mathsf{E}(X^2) = (-1)^\tau\, \mathsf{Var}(X) \end{aligned}$$

(d) If $\{X_t\}$ is stationary then it follows from (c):

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \frac{(-1)^\tau\, \mathsf{Var}(X)}{\mathsf{Var}(X)} = (-1)^\tau.$$

Exercise 11.2 (Empirical Analysis). *Perform an empirical analysis using the data on DAX and Dow Jones index from the period Jan. 1, 1997 to Dec. 30, 2004.*

(a) Display a time plot of the given indices data, its returns and log returns.
(b) Calculate mean, skewness, kurtosis, autocorrelation of the first order, autocorrelation of squared returns, and autocorrelation of absolute returns for the given data.

(a) The time series plot for the DAX and Dow Jones indices are represented by Fig. 11.2. One can observe from the figure that stock markets have fallen since September 11, 2001. However, shortly after the catastrophe, the indices values experienced a moderate increase until they climbed up to their original values.

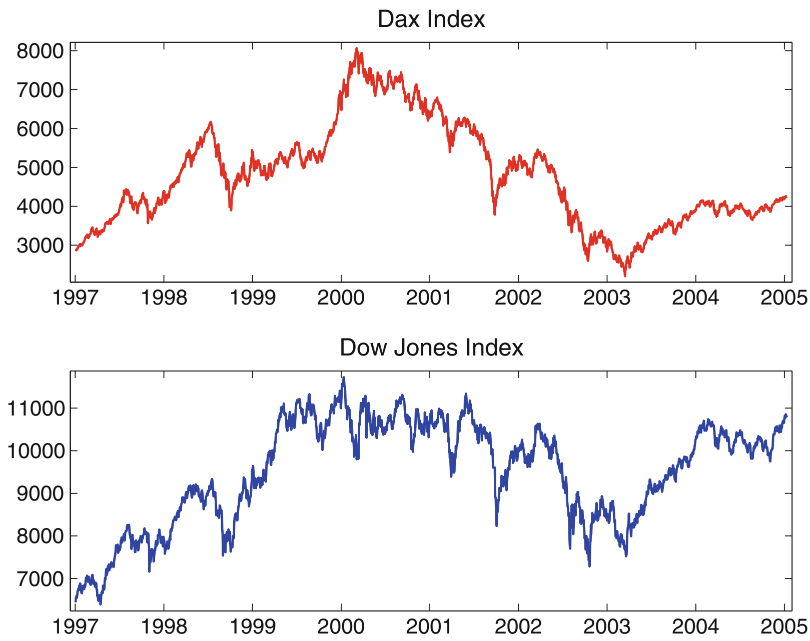


Fig. 11.2 Time series plot for DAX index (*upper panel*) and Dow Jones index (*lower panel*) from the period Jan. 1, 1997 to Dec. 30, 2004. `SFStimeseries`

The returns and the log returns are represented by Figs. 11.3 and 11.4 respectively.

(b) Denote μ mean, S skewness, Kurt kurtosis, $\rho_1(r_t)$ autocorrelation of the first order, $\rho_1(r_t^2)$ autocorrelation of squared returns, $\rho_1(\|r_t\|)$ autocorrelation of absolute returns. Table 11.1 summarises the results.

Exercise 11.3 (Distribution of Returns and Test of Normality). *Consider the data on the DAX and Dow Jones index from the Exercise 11.2. Which empirical distribution do the returns follow? Are they normally distributed? Perform an appropriate test of normality.*

From Table 11.1 one can observe that the kurtosis is larger than 3, i.e. the distribution is leptokurtic. In addition, the skewness is smaller than zero, i.e. right side tilted. Figure 11.5 represents density functions of DAX and Dow Jones in comparison to normal density, estimated nonparametrically with Gaussian kernel.

We use the Bera-Jarque test for normality. We test the H_0 hypothesis of normality against an alternative H_1 to establish that the data is not normal distributed. The test statistics of the Bera-Jarque test (see Franke et al. (2011)) is given by:

$$BJ = n \left\{ \frac{\widehat{S^2}}{6} + \frac{(\widehat{Kurt} - 3)^2}{24} \right\} \xrightarrow{\mathcal{L}} \chi_2^2 \text{ under } H_0,$$

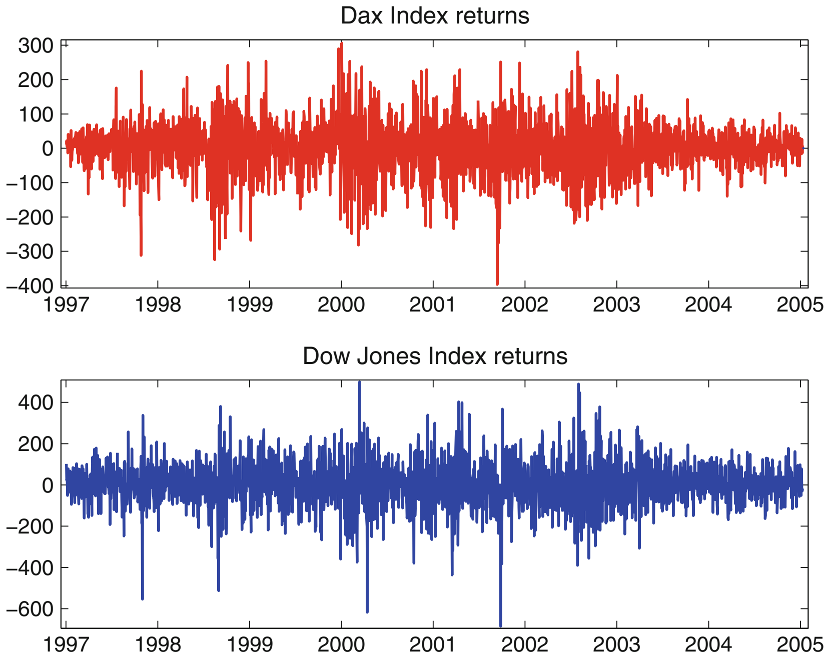


Fig. 11.3 Returns of DAX (*upper panel*) and Dow Jones (*lower panel*) from the period Jan. 1, 1997 to Dec. 30, 2004. `SFStimeseries`

which means that under the H_0 hypothesis the data is normal distributed. Calculating the Bera-Jarque test statistics for DAX and Dow Jones leads us to 357.682 and 999.89 respectively. Comparing the values of the test statistics with a 5 % – critical value of the χ^2_2 distribution, which is 5.99, we reject the hypothesis of normality for DAX and Dow Jones returns. One can obtain the same results comparing the $P-$ value, which is 0.00 for both DAX and Dow Jones, at level $\alpha = 0.05$.

Exercise 11.4 (Proof of Stylized Facts). *According to the stylized facts, the autocorrelation of first order is close to zero for all stock returns; the autocorrelation of squared and absolute returns are positive and different from zero. In addition, small (positive or negative) returns are followed by small (positive or negative) returns and large returns are followed by large returns. Can you prove these facts by applying them to your data? Plot the autocorrelation function (ACF) for returns, absolute returns and for squared log returns from the DAX and Dow Jones data from Exercise 11.1.*

The sample autocorrelation function (ACF) for returns, absolute returns and squared log returns together with the 95 % confidence band are represented by Figs. 11.6–11.8 respectively. In fact, we can see from the plots that the autocorrelation of first order is close to zero and that small returns are followed

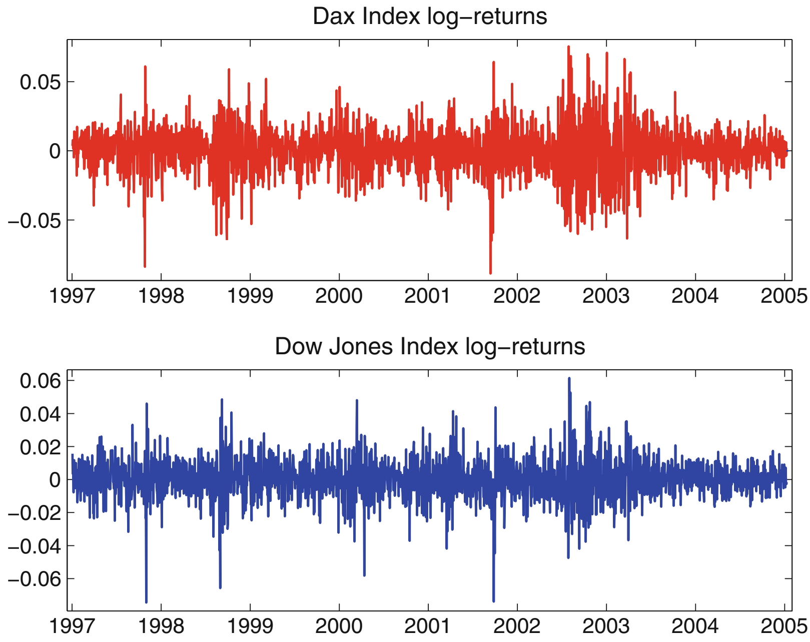


Fig. 11.4 Log-returns of DAX (*upper panel*) and Dow Jones (*lower panel*) from the period Jan. 1, 1997 to Dec. 30, 2004. SFStimeseries

Table 11.1 Descriptive statistics for the DAX index (*upper line*) and the Dow Jones index (*lower line*)

μ	S	Kurt	$\rho_1(r_t)$	$\rho_1(r_t^2)$	$\rho_1(\|r_t\|)$
0.000196	−0.19558	5.02	−0.0094	0.1875	0.185
0.000255	−0.20065	6.42	−0.0123	0.1351	0.116

by small returns, large returns are followed by large returns. In addition, one can observe that the autocorrelation of squared returns is positive and different from zero.

Exercise 11.5 (Augmented Dickey-Fuller Test). *Use the data of DAX and Dow Jones indices from Exercise 11.1. Apply the Augmented Dickey-Fuller test (ADF) of stationarity to the*

(a) Raw data, i.e. $I(0)$
(b) Log returns, i.e. $I(1)$.

First, consider a regression model without linear time trend and then with linear time trend. Can we reject the hypothesis of trend stationarity in both cases?

We consider the autoregressive process of first order:

$$AR(1) : X_t = c + \alpha X_{t-1} + \varepsilon_t.$$

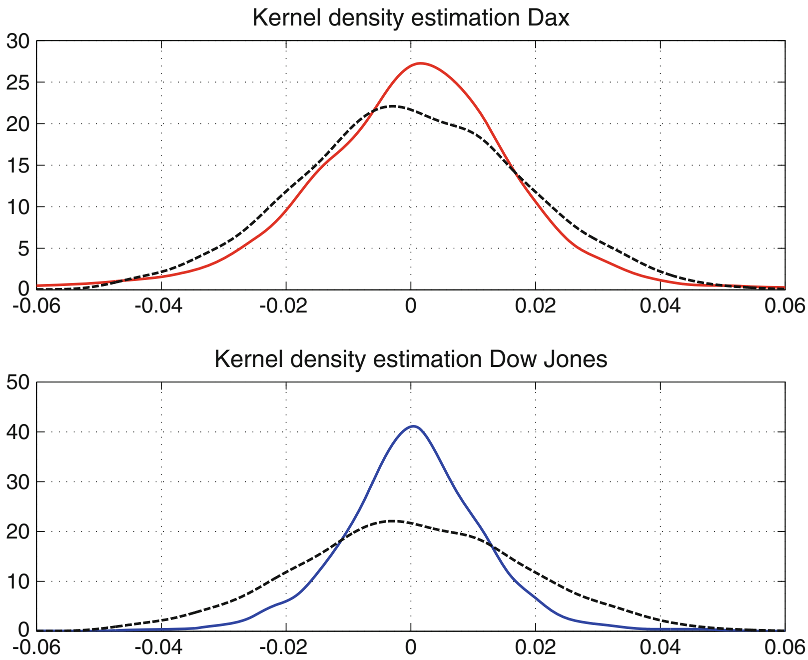


Fig. 11.5 Density functions of DAX (*upper panel*) and Dow Jones (*lower panel*) and the normal density (*dashed line*), estimated nonparametrically with Gaussian kernel. `SFSdaxdowkernel`

We know that if $|\alpha| < 1$, then the process X_t is stationary and for $|\alpha| = 1$, the process X_t is a random walk, i.e. non stationary, see Franke et al. (2011). We apply the Augmented Dickey-Fuller test (ADF) to test $\alpha = 1$. First, we test a regression model without linear time trend:

$$\varDelta X_t = c + (\alpha - 1)X_{t-1} + \sum_{i=1}^{p} \alpha_i \varDelta X_{t-i} + \varepsilon_t$$

and then a regression model with a linear time trend:

$$\varDelta X_t = c + \mu t + (\alpha - 1)X_{t-1} + \sum_{i=1}^{p} \alpha_i \varDelta.$$

We test the hypothesis H_0 of non stationarity (i.e. $\alpha = 1$) against the alternative H_1 hypothesis ($\alpha \neq 1$). The test statistics of the Augmented Dickey-Fuller test is given by

$$\widehat{t}_n = \frac{1 - \widehat{\alpha}}{\sqrt{\widehat{\sigma}^2} \sum_{t=2}^{n} X_{t-1}^2} \xrightarrow{\mathcal{L}} \frac{W^2(1) - 1}{2\left\{\int_0^1 W^2(u)du\right\}^{1/2}}.$$

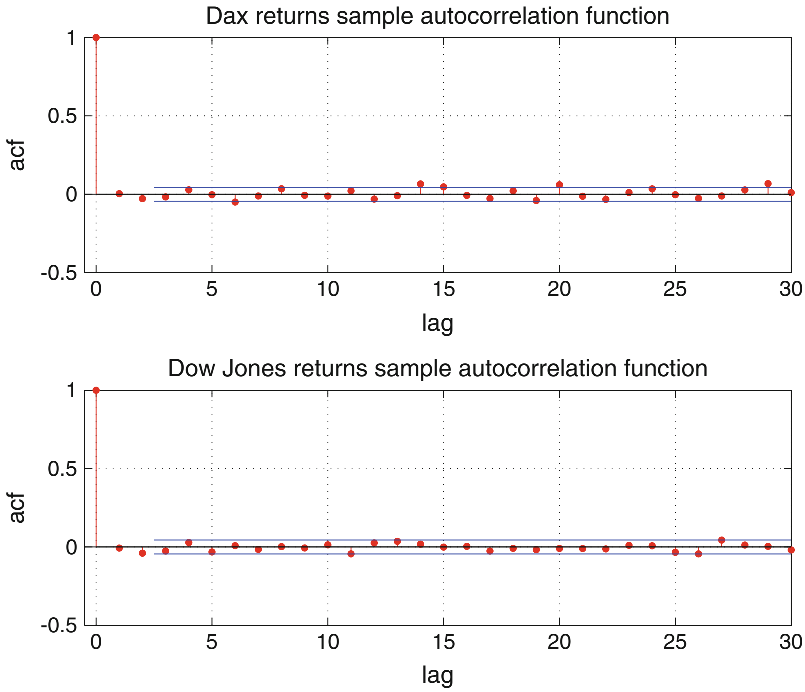


Fig. 11.6 Autocorrelation function for the DAX returns (*upper panel*) and Dow Jones returns (*lower panel*). `SFStimeseries`

The hypothesis H_0 will be rejected if $\widehat{t}_n$ is smaller than the critical value.

(a) We apply the ADF to the DAX and Dow Jones raw data, i.e. $I(0)$. In the regression model without linear time trend, the values of the test statistics correspond to -1.7094 for the DAX and -2.6058 for the Dow Jones and thus, are smaller than a 5 % critical value, which corresponds to -2.86. Analogically, we calculate the values of the ADF test statistics for the regression model with a linear time trend. It leads us to the values of -2.1117 for the DAX and -2.5719 for the Dow Jones, which are below the 5 % critical value corresponding to -3.41. Hence, we can not reject H_0 in both cases and thus, the process is not trend stationary.

(b) We now apply the test to the DAX and Dow Jones log returns, i.e. $I(1)$. In the regression model without linear time trend, the values of the test statistics correspond to -20.555 for the DAX and -21.068 for the Dow Jones and thus, exceed the 5 % critical value. The same situation occurs for the regression model with a linear time trend, where the values of the test statistics correspond to -20.597 for the DAX and -21.08 for the Dow Jones. In both cases, one rejects H_0, i.e. the process is trend stationary.

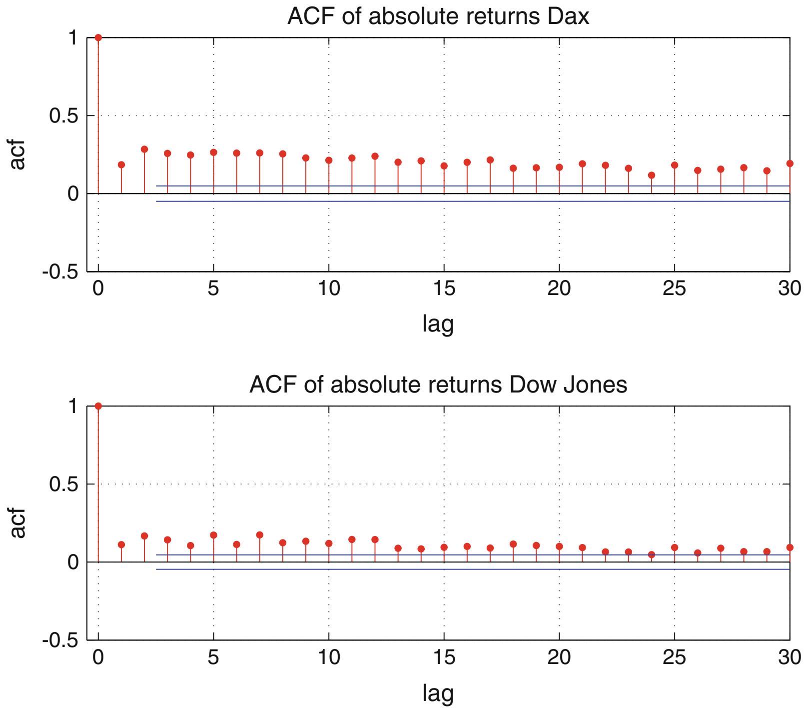


Fig. 11.7 Autocorrelation function for the DAX absolute returns (*upper panel*) and Dow Jones absolute returns (*lower panel*). `SFStimeseries`

Exercise 11.6 (KPSS Test of Stationarity). *Use the data on DAX and Dow Jones indices from Exercise 11.2. Apply the KPSS test of stationarity to the*

(a) Raw data, i.e. $I(0)$
(b) Log returns, i.e. $I(1)$

First, consider a regression model with constant μ *and then with linear time trend. Can we reject the hypothesis of trend stationarity in both cases?*

Firstly, we consider a regression model with a constant μ:

$$X_t = c + k \sum_{i=1}^{t} \xi_i + \eta_t$$

and then with linear time trend:

$$X_t = c + \mu t + k \sum_{i=1}^{t} \xi_i + \eta_t.$$

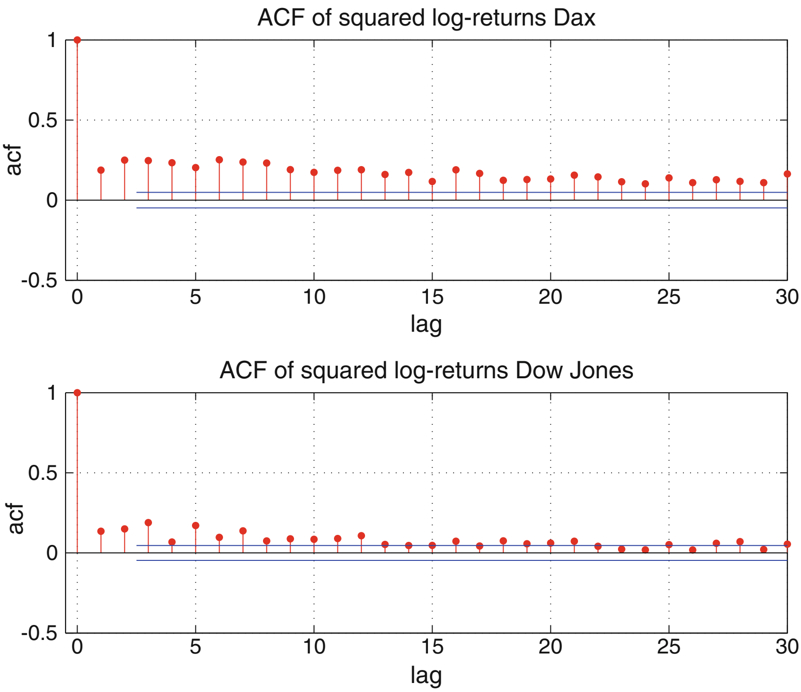


Fig. 11.8 Autocorrelation function for the DAX squared log-returns (*upper panel*) and Dow Jones squared log-returns (*lower panel*). `SFStimeseries`

We test the hypothesis H_0 of trend stationary (i.e. $k = 0$) against the alternative: $k \neq 0$, i.e. non stationarity. The test statistics, see Franke et al. (2011), is given by

$$KPSS_T = \frac{\sum_{t=1}^{n} S_t^2}{n^2 \widehat{\omega}_T^2}.$$

We reject H_0 if $KPSS_T$ is larger than the critical value.

(a) We apply the KPSS test to the DAX and Dow Jones raw data, i.e. $I(0)$. In the regression model with a constant μ, the values of the test statistics correspond to 51.414 for the DAX and 47.441 for the Dow Jones and thus exceed the 5 % critical value, corresponding to 0.463. Analogically, we calculate the values of the KPSS test statistics for the regression model with linear time trend. It leads us to the values of 34.138 for the DAX and 29.163 for the Dow Jones, which exceed the 5 % critical value corresponding to -3.41. Hence, we reject H_0, i.e. the process is not trend stationary.

(b) We now apply the test to the DAX and Dow Jones log returns, i.e. $I(1)$. In the regression model with a constant μ, the values of the test statistics correspond to 0.29653 for the DAX and 0.13115 for the Dow Jones which are smaller than the 5 % critical value. We observe the same situation for the regression model with linear time trend, where the values of the test statistics correspond to 0.131 for the DAX and 0.064037 for the Dow Jones. In both cases, one can not reject H_0, i.e. the process is trend stationary.

Exercise 11.7 (ADF and KPSS Test). *What is the difference between ADF and KPSS test?*

The ADF tests in favour of nonstationarity, i.e. against trend stationarity ($H_0 : \alpha=1$). We reject H_0 if the value of the ADF test statistics is smaller than the critical value. The KPSS tests in favour of stationarity, i.e. against nonstationarity ($H_0 : \alpha < 1$). We reject H_0 if the value of the KPSS test statistics is larger than the critical value.

Chapter 12
ARIMA Time Series Models

अरीमा समय श्रृंखला माङ्ल

समय किसी की प्रतीक्षा नही करता।
Time does not wait for anyone.

The autoregressive moving average (ARMA) model defined as

$$X_t = \nu + \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q} + \varepsilon_t,$$

deals with linear time series. The time series should be a covariance stationary process. It consists of two parts, an autoregressive (AR) part of order p and a moving average (MA) part of order q. When an ARMA model is not stationary, the methods of analyzing stationary time series cannot be used directly. In order to handle those processes within the framework of the classical time series analysis, we must first form the differences to get a stationary process. The autoregressive integrated moving average (ARIMA) models are an extension of ARMA processes by the integrated (I) part. Sometimes ARIMA models are referred to as ARIMA(p, d, q) whereas p and q denote the order of an autoregressive (AR) respective a moving average (MA) part and d describes the integrated (I) part.

Exercise 12.1 (ARIMA Model). *Which condition do time series have to fulfill in order to be fitted by the ARIMA model? And what does the letter "I" in the word "ARIMA" mean?*

To be fitted by the ARIMA model, the underlying time series should be covariance stationary processes. "I" in ARIMA stands for "Integrated". We say that the process X_t is integrated of order d, $I(d)$, when $(1 - L)^{d-1} X_t$ is non-stationary and

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_12,

$(1 - L)^d X_t$ is stationary, see Franke et al. (2011). The integrated part of the model determines whether the observed values are modelled directly, or whether the differences between the observations are modelled instead. If the order of the integrated part $d = 0$, the observed values are modelled directly. If $d = 1$ or $d = 2$ (the first and the second order respectively) then the differences between the observed values are modelled.

Exercise 12.2 (Autocorrelation Function). *Suppose that the stationary process X_t has an autocovariance function given by γ_τ. Find the autocorrelation function (in terms of γ_τ) of the (stationary) process Y_t defined as $Y_t = X_t - X_{t-1}$.*

If the process Y_t is stationary, the autocorrelation function is given by

$$\rho_\tau(Y_t) = \frac{\gamma_\tau(Y_t)}{\gamma_0(Y_t)}$$

where $\gamma_\tau(Y_t)$ and $\gamma_0(Y_t)$ denote the autocovariance function and the variance of Y_t respectively. The autocovariance function of Y_t is given by:

$$\begin{aligned}
\gamma_\tau(Y_t) &= \mathsf{Cov}(Y_t, Y_{t-\tau}) = \mathsf{Cov}(X_t - X_{t-1}, X_{t-\tau} - X_{t-\tau-1}) \\
&= \mathsf{Cov}(X_t, X_{t-\tau}) - \mathsf{Cov}(X_t, X_{t-\tau-1}) - \mathsf{Cov}(X_{t-1}, X_{t-\tau}) + \mathsf{Cov}(X_{t-1}, X_{t-\tau-1}) \\
&= \gamma_\tau - \gamma_{\tau+1} - \gamma_{\tau-1} + \gamma_\tau \\
&= 2\gamma_\tau - \gamma_{\tau-1} - \gamma_{\tau+1}
\end{aligned}$$

and $\gamma_0(Y_t) = \mathsf{Var}(Y_t) = 2\gamma_0 - 2\gamma_1$.

For the autocorrelation of Y_t we therefore obtain:

$$\rho_\tau(Y_t) = \frac{2\gamma_\tau - \gamma_{\tau-1} - \gamma_{\tau+1}}{2\gamma_0 - 2\gamma_1}.$$

Exercise 12.3 (Autocorrelation Function of MA(1) Process). *Calculate the autocorrelation function (ACF) of the MA(1) process $X_t = -0.5\varepsilon_{t-1} + \varepsilon_t$.*

For the MA(q) process with $\beta_0 = 1$ and $\mathsf{E}(X_t) = 0$, i.e.,

$$X_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q}$$

the covariance structure is given by

$$\gamma_\tau = \mathsf{Cov}(X_t, X_{t+\tau}) = \sum_{i=0}^{q-\tau} \beta_i \beta_{i+\tau} \sigma^2, \quad |\tau| \le q.$$

For the autocorrelation function we therefore obtain

$$\rho_0 = \frac{\sum_{i=0}^{1-0} \beta_i \beta_{i+0}}{\sum_{i=0}^{1} \beta_i^2} = \frac{\sum_{i=0}^{1} \beta_i^2}{\sum_{i=0}^{1} \beta_i^2} = 1$$

$$\rho_1 = \frac{\sum_{i=0}^{1-1} \beta_i \beta_{i+1}}{\sum_{i=0}^{1} \beta_i^2} = \frac{\sum_{i=0}^{0} \beta_i \beta_{i+1}}{\sum_{i=0}^{1} \beta_i^2} = \frac{\beta_0 \beta_1}{\beta_0^2 + \beta_1^2} = \frac{-0.5}{1 + (-0.5)^2} = -0.4$$

and $\rho_\tau = 0$ for $\tau > 1$.

Exercise 12.4 (Autocorrelation Function of MA(2) Process). *Find the autocorrelation function of the second order moving average process MA(2) defined as*

$$X_t = \varepsilon_t + 0.5\varepsilon_{t-1} - 0.2\varepsilon_{t-2}$$

where ε_t denotes white noise.

The covariance function for the MA(q) process with $\beta_0 = 1$ and $\mathsf{E}(X_t) = 0$ is given by

$$\gamma_\tau = \mathsf{Cov}(X_t, X_{t+\tau}) = \sum_{i=0}^{q-\tau} \beta_i \beta_{i+\tau} \sigma^2, \quad |\tau| \le q.$$

Therefore it holds:

$$\begin{aligned}
\gamma_0 &= \mathsf{Cov}(X_t, X_t) = \sigma^2(\beta_0\beta_0 + \beta_1\beta_1 + \beta_2\beta_2) \\
&= \sigma^2(1 + 0.25 + 0.04) = 1.29\sigma^2 \\
\gamma_1 &= \mathsf{Cov}(X_t, X_{t+1}) = \sigma^2(\beta_0\beta_1 + \beta_1\beta_2) \\
&= \sigma^2\{1 \cdot 0.5 + 0.5 \cdot (-0.2)\} = 0.4\sigma^2 \\
\gamma_2 &= \mathsf{Cov}(X_t, X_{t+2}) = \sigma^2(\beta_0\beta_2) \\
&= \sigma^2\{1 \cdot (-0.2)\} = -0.2\sigma^2.
\end{aligned}$$

For the autocorrelation function we have:

$$\begin{aligned}
\rho_0 &= 1 \\
\rho_1 &= \frac{\gamma_1}{\gamma_0} = \frac{0.4\sigma^2}{1.29\sigma^2} = 0.31008 \\
\rho_2 &= \frac{\gamma_2}{\gamma_0} = \frac{-0.2\sigma^2}{1.29\sigma^2} = -0.15504 \\
\rho_k &= 0 \quad \text{for} \quad k \ge 3.
\end{aligned}$$

Exercise 12.5 (Autocorrelation Function of MA(m) Process). *Let*

$$X_t = \sum_{k=0}^{m} \frac{1}{m+1} \varepsilon_{t-k}$$

be the m-th order moving average process MA(m). Show that the autocorrelation function ACF of this process is given by

$$\rho_\tau = \left\{ \begin{array}{lll} (m+1-k)/(m+1) & \textit{if} & k = 0, 1, \ldots, m \\ 0 & \textit{if} & k > m \end{array} \right\}.$$

For the MA(m) process

$$X_t = \sum_{k=0}^{m} \frac{1}{m+1} \varepsilon_{t-k}$$

we have:

$$\gamma_0 = \mathsf{Cov}(X_t, X_t) = \sum_{i=0}^{m} \beta_i \beta_i \sigma^2$$

$$= \sum_{i=0}^{m} \frac{1}{(m+1)^2} \sigma^2 = (m+1) \frac{1}{(m+1)^2} \sigma^2 = \frac{\sigma^2}{m+1}$$

$$\gamma_1 = \mathsf{Cov}(X_t, X_{t-1}) = \sum_{i=0}^{m-1} \beta_i \beta_{i+1} \sigma^2$$

$$= \sum_{i=0}^{m-1} \frac{1}{(m+1)^2} \sigma^2 = m \frac{1}{(m+1)^2} \sigma^2$$

$$\gamma_2 = \mathsf{Cov}(X_t, X_{t-2}) = \sum_{i=0}^{m-2} \beta_i \beta_{i+2} \sigma^2$$

$$= \sum_{i=0}^{m-2} \frac{1}{(m+1)^2} \sigma^2 = (m-1) \frac{1}{(m+1)^2} \sigma^2$$

$$\gamma_k = \mathsf{Cov}(X_t, X_{t-k}) = \sum_{i=0}^{m-k} \beta_i \beta_{i+1} \sigma^2$$

$$= \sum_{i=0}^{m-k} \frac{1}{(m+1)^k} \sigma^2 = (m+1-k) \frac{1}{(m+1)^2} \sigma^2$$

The autocorrelation function of this process is therefore given by

$$\rho_0 = 1$$
$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{m}{(m+1)^2}\sigma^2 \cdot \frac{m+1}{\sigma^2} = \frac{m}{m+1}$$
$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{m-1}{(m+1)^2}\sigma^2 \cdot \frac{m+1}{\sigma^2} = \frac{m-1}{m+1}$$
$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{m+1-k}{(m+1)^2}\sigma^2 \cdot \frac{m+1}{\sigma^2} = \frac{m+1-k}{m+1}.$$

For $k > m$ we have $\gamma_k = 0$ and thus, $\rho_k = 0$.

Exercise 12.6 (Stationarity, Invertibility and the Shift Operator L).

(a) What is meant by saying that a linear process is stationary? How can we evaluate whether a process is stationary?
(b) What is meant by saying that a linear process is invertible? How can we evaluate whether a process is invertible?
(c) For each of the following models express the model in terms of the shift operator L acting on ε_t and determine whether the model is stationary or/and invertible or not.

(i) $X_t = 0.2X_{t-1} + \varepsilon_t$
(ii) $X_t = \varepsilon_t - 1.5\varepsilon_t + 0.3\varepsilon_{t-2}$
(iii) $X_t = 0.4X_{t-1} + \varepsilon_t - 1.5\varepsilon_{t-1} + 0.3\varepsilon_{t-2}$

(a) A linear process is stationary if it can be written in a moving average form $X_t = \beta(L)\varepsilon_t$ with $\beta(L) = 1 + \beta_1 L + \ldots + \beta_q L^q$.
The AR(p) process $X_t = \nu + \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p}$ is stationary if all roots z_i of the characteristic equation $\alpha(z) = 1 - \alpha_1 z - \ldots - \alpha_q z^q$ lie outside of the complex unit circle.

(b) A linear process is invertible if it can be written in an autoregressive form $\alpha(L)X_t = \nu + \varepsilon_t$ with $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_q L^q$.
The MA(q) process is invertible if all roots z_i of the characteristic equation $\beta(z) = 1 + \beta_1 z + \ldots + \beta_q z^q$ lie outside of the complex unit circle. In this case holds: $\beta(L)\beta^{-1}(L) = 1$.

(i) The process
$$X_t = 0.2X_{t-1} + \varepsilon_t$$
can be written as:
$$(1 - 0.2L)X_t = \varepsilon_t.$$
The model is stationary if for the root of the equation
$$1 - 0.2z = 0$$
holds: $|z| > 1$. Since $z = 1/0.2 = 5 > 1$, the process is stationary.

The process is invertible, since it has an autoregressive representation.

(ii) The process

$$X_t = \varepsilon_t - 1.5\varepsilon_t + 0.3\varepsilon_{t-2}$$

can be written as:

$$X_t = (1 - 1.5L + 0.3L^2)\varepsilon_t.$$

The model is invertible if for the roots $|z_i|$ of the equation

$$1 - 1.5z + 0.3z^2 = 0$$

holds: $|z_i| > 1$. Since $z_1 = 4.2 > 1$ and $z_2 = 0.8 < 1$, the process is not invertible.

The process is stationary since it has a moving average representation.

(iii) The process

$$X_t = 0.4X_{t-1} + \varepsilon_t - 1.5\varepsilon_{t-1} + 0.3\varepsilon_{t-2}$$

can be written as

$$(1 - 0.4L)X_t = (1 - 1.5L + 0.3L^2)\varepsilon_t.$$

The model is stationary since for the root $|z|$ of the equation

$$1 - 0.4z = 0$$

holds: $z = 2.5 > 1$, the process is stationary. The process is invertible if for the roots $|z_i|$ of the equation

$$1 - 1.5z + 0.3z^2 = 0$$

holds: $|z_i| > 1$. From (ii) follows that the process is not invertible.

Exercise 12.7 (Partial Autocorrelation). *Calculate the partial autocorrelations of first and second order of the AR(1) process $X_t = 0.5X_{t-1} + \varepsilon_t$ by using Yule-Walker equations.*

From Yule-Walker equations we know:

$$\alpha_1 + \rho_1\alpha_2 = \rho_1, \text{ hence } \rho_1 = 0.5$$

$$\rho_1\alpha_1 + \alpha_2 = \rho_2, \text{ hence } \rho_2 = 0.25$$

We therefore obtain, see Franke et al. (2011):

$$\phi_{11} = \frac{|P_1^*|}{|P_1|} = \frac{\rho_1}{1} = \rho_1 = 0.5,$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{0.25 - 0.5^2}{1 - 0.5^2} = 0.$$

Exercise 12.8 (Estimating ARIMA Model Parameters). *Which methods could we use to estimate the parameters in an ARIMA model? What are the advantage and drawback of each of them?*

Estimator	Advantage	Drawback
Yule-Walker	Simple to estimate	Asymptotically inefficient
Least squares	Asymptotically efficient	Solution only with iterative numerical algorithms
And asymptotic normal distributed	Only under some technical assumptions	

Exercise 12.9 (Adequacy of fitted ARIMA Model). *Could you give some statistical tests to assess the adequacy of a fitted ARIMA model?*

First check whether the coefficients of the ACF and PACF are equal to zero; then check whether the Portmanteau statistics take small values.

Exercise 12.10 (Characteristics of MA(1) process). *What characteristics would one expect of a realization of the MA(1) process* $Y_t = 1 + \varepsilon_t + 0.8\varepsilon_{t-1}$? *How would these characteristics differ from the those of a realization of the process?* $Y_t' = 1 + \varepsilon_t' - 0.8\varepsilon_{t-1}'$

(i) The correlation does not extend more than one period out, so that the realization appears very "noisy".
(ii) Y_t tends to be positively correlated with adjacent values; e.g., a positive value is more likely to be preceded and followed by a positive value than by a negative value.
(iii) A realization of the process $Y_t' = 1 + \varepsilon_t' - 0.8\varepsilon_{t-1}'$ would show negative correlations between adjacent values, so that a positive value of Y_t' would be more likely to be followed by a negative value.

Exercise 12.11 (Covariance and Autocorrelation for MA(3) Process). *Calculate the covariances* γ_k *for MA(3), the moving average of order 3. Determine the autocorrelation function for this process. Plot the autocorrelation function for the MA(3) process:*

$$Y_t = 1 + \varepsilon_t + 0.8\varepsilon_{t-1} - 0.5\varepsilon_{t-2} + 0.3\varepsilon_{t-3}.$$

The variance and covariances are given by:

$$\gamma_0 = \sigma_\varepsilon^2(1 + \theta_1^2 + \theta_2^3 + \theta_3^2)$$
$$\gamma_1 = \sigma_\varepsilon^2(-\theta_1 + \theta_2\theta_1 + \theta_3\theta_2)$$
$$\gamma_2 = \sigma_\varepsilon^2(-\theta_2 + \theta_3\theta_1)$$
$$\gamma_3 = -\theta_3\sigma_\varepsilon^2$$
$$\gamma_k = 0,\ k > 3$$

where $\theta_1 = 0.8$, $\theta_2 = -0.5$, $\theta_3 = 0.3$. The autocorrelation plot is represented by the Fig. 12.1.

Exercise 12.12 (Autocorrelation Function for ARMA (2,1) Process). *Derive the autocorrelation function for the ARMA(2,1) process:*

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

that is, determine ρ_1*,* ρ_2*, etc., in term of* ϕ_1*,* ϕ_2*, and* θ_1*.*

According to the Yule-Walker Equations, the autocorrelations are:

$$\rho_1 = \frac{\phi_1}{1-\phi_2} - \frac{\theta_1(1-\phi_1^2\phi_2 - \phi_1^2 - \phi_2)}{(1-\phi_2)^2(1-2\phi_1\theta_1+\theta_2^2) - 2\phi_1\phi_2\theta_1(1-\phi_2)}$$
$$\rho_2 = \phi_2 + \phi_1\rho_1$$
$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1.$$

Exercise 12.13 (Forecasting). *Derive expressions for the one-, two-, three-period forecast,* $\hat{Y}_t(1)$*,* $\hat{Y}_t(2)$*, and* $\hat{Y}_t(3)$*, for the second-order autoregressive process AR(2). What are the variances of the errors for these forecasts?*

The one-, two-, three-period forecasts for the AR(2) process are the following:

$$\hat{Y}_t(1) = \phi_1 Y_t + \phi_2 Y_{t-1} + \delta$$
$$\hat{Y}_t(2) = \phi_1 \hat{Y}_t(1) + \phi_2 Y_t + \delta = (\phi_1^2 + \phi_2)Y_t + \phi_1\phi_2 Y_{t-1} + (1+\phi_1)\delta$$
$$\hat{Y}_t(3) = (\phi_1^3 + 2\phi_1\phi_2)Y_t + (\phi_1^2\phi_2 + \phi_2^2)Y_{t-1} + (1 + \phi_1 + \phi_1^2 + \phi_2)\delta$$

These forecasts have error variances:

$$\mathsf{E}[\varepsilon_t^2(1)] = \sigma_\varepsilon^2$$
$$\mathsf{E}[\varepsilon_t^2(2)] = (1+\phi_1^2)\sigma_\varepsilon^2$$
$$\mathsf{E}[\varepsilon_t^2(3)] = \{1 + \phi_1^4 + \phi_1^2(1+2\phi_2) + \phi_2^2\}\sigma_\varepsilon^2$$

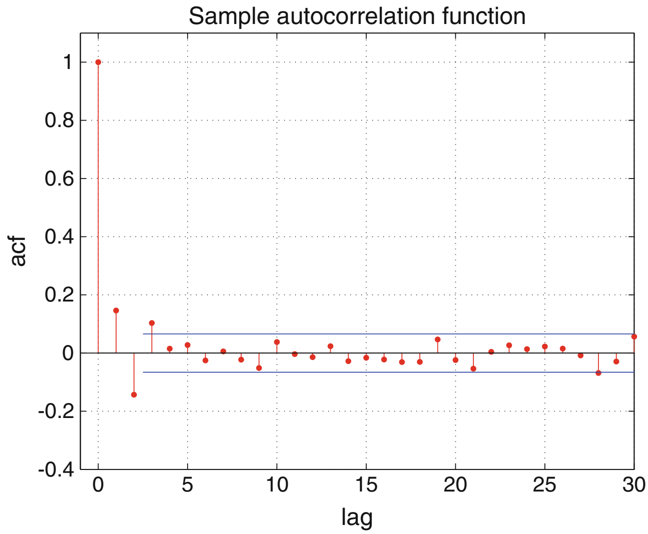


Fig. 12.1 The autocorrelation function for the MA(3) process: $Y_t = 1 + \varepsilon_t + 0.8\varepsilon_{t-1} - 0.5\varepsilon_{t-2} + 0.3\varepsilon_{t-3}$ SFSacfMA3

Exercise 12.14 (Diagnostic Test). *Suppose an ARMA(0,2) model has been estimated from a time series generated by an ARMA(1,2) process. How would the diagnostic test indicate that the model has been misspecified?*

This misspecification will result in residuals that are autocorrelated. The diagnostic Portmanteau Q is likely to be e.g. above 90 % on the χ^2 distribution.

Exercise 12.15 (Covariance Stationarity). *Which of the following processes are covariance stationary? Explain your answer for every process.*

(a) $X_t = \varepsilon_t - \varepsilon_{t-1} + 2\varepsilon_{t-2}$
(b) $X_t = 2 + \frac{1}{2}X_{t-1} + \varepsilon_t$
(c) $X_t = 4 - 1.3X_{t-1} + 0.8X_{t-2} + \varepsilon_t$
(d) $X_t = 4 - 1.3X_{t-1} - 0.8X_{t-2} + \varepsilon_t$

(a) This process is stationary because all moving average processes are stationary as a linear combination of a stationary white noise processes.
(b) The stationarity of this process follows from a general statement:
An AR(1) process $X_t = \alpha_0 + \alpha_1 X_{t-1} + \varepsilon_t$ is stationary if $|\alpha_1| < 1$ or if the root z of the characteristic equation $\alpha(z) = 1 - 1/2z = 0$ lie outside of the complex unit circle. For this process holds: $|\alpha_1| = 1/2 < 1$ or $|z| = 2 > 1$, and thus the process is stationary.
(c) This process is not stationary since the coefficients do not lie in the "triangle of stationarity": $\alpha_1 + \alpha_2 < 1$ and $\alpha_2 - \alpha_1 < 1$ or the roots $|z_i|$ of the characteristic equation $\alpha(z) = 1 + 1.3z - 0.8z^2 = 0$ do not lie outside of the complex unit circle ($|z_1| = 2.1946 > 1$ but $|z_2| = 0.56958 < 1$).

(d) This process is stationary because the coefficients lie in the "triangle of stationarity": $\alpha_1 + \alpha_2 = -2.1 < 1$ and $\alpha_2 - \alpha_1 = 0.5 < 1$.

Exercise 12.16 (Autocorrelation Function). *Find the autocorrelation function for the following processes:*

(a) A white noise process with $\mathsf{E}(X_t) = 0$, $\mathsf{Var}(X_t) = \sigma^2\ \forall t$
(b) $X_t = \varepsilon_t - \varepsilon_{t-1}$
(c) For the MA(1) process defined as $X_t = \varepsilon_t - \theta_1\varepsilon_{t-1}$. *Show that you cannot identify an MA(1) process uniquely from the autocorrelation by comparing the results using* θ_1 *with those if you replace* θ_1 *by* θ_1^{-1}.

(a) The autocovariance function of a white noise process is defined by

$$\gamma_\tau = \begin{cases} \sigma^2 & \text{if} \quad \tau = 0 \\ 0 & \text{if} \quad \tau \neq 0 \end{cases}$$

Since white noise is stationary, the autocorrelation function is given by

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \begin{cases} 1 & \text{if} \quad \tau = 0 \\ 0 & \text{if} \quad \tau \neq 0 \end{cases}$$

(b) For the process $X_t = \varepsilon_t - \varepsilon_{t-1}$ we have

$$\begin{aligned} \mathsf{E}(X_t) &= 0 \\ \gamma(t,\tau) &= \mathsf{E}\{(\varepsilon_t - \varepsilon_{t-1})(\varepsilon_{t+\tau} - \varepsilon_{t+\tau-1})\} \\ &= \mathsf{E}(\varepsilon_t\varepsilon_{t+\tau}) - \mathsf{E}(\varepsilon_t\varepsilon_{t+\tau-1}) - \mathsf{E}(\varepsilon_{t-1}\varepsilon_{t+\tau}) + \mathsf{E}(\varepsilon_{t-1}\varepsilon_{t+\tau-1}) \\ &= \begin{cases} 2\sigma^2 & \text{if} \quad \tau = 0 \\ -\sigma^2 & \text{if} \quad \tau = \pm 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Since the process $X_t = \varepsilon_t - \varepsilon_{t-1}$ is a linear combination of a stationary white noise processes and thus, stationary, the autocorrelation function of X_t is given by

$$\rho_\tau = \begin{cases} 1 & \text{if} \quad \tau = 0 \\ -1/2 & \text{if} \quad \tau = \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

(c) For the MA(1) process $X_t = \varepsilon_t - \theta_1\varepsilon_{t-1}$ we have:

$$\begin{aligned}
\mathsf{E}(X_t) &= 0\\
\gamma(t,\tau) &= \mathsf{E}\{(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_{t+\tau} - \theta_1\varepsilon_{t+\tau-1})\}\\
&= \mathsf{E}(\varepsilon_t\varepsilon_{t+\tau}) - \theta_1\,\mathsf{E}(\varepsilon_t\varepsilon_{t+\tau-1}) - \theta_1\,\mathsf{E}(\varepsilon_{t-1}\varepsilon_{t+\tau}) + \theta_1^2\,\mathsf{E}(\varepsilon_{t-1}\varepsilon_{t+\tau-1})\\
&= \begin{cases} \sigma^2 + \theta_1^2\sigma^2 = \sigma^2(1+\theta_1^2) & \text{if} \quad \tau = 0\\ -\theta_1\sigma^2 & \text{if} \quad \tau = \pm 1\\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Since X_t is a stationary process, the autocorrelation function is given by

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \begin{cases} 1 & \text{if} \quad \tau = 0\\ -\frac{\theta_1}{1+\theta_1^2} & \text{if} \quad \tau = \pm 1\\ 0 & \text{otherwise} \end{cases}$$

If we replace θ_1 by θ_1^{-1}, the model becomes

$$X_t = \varepsilon_t - \theta_1^{-1}\varepsilon_{t-1}$$

with

$$\gamma(t,\tau) = \begin{cases} \sigma^2(1+\frac{1}{\theta_1^2}) & \text{if} \quad \tau = 0\\ -\frac{1}{\theta_1}\sigma^2 & \text{if} \quad \tau = \pm 1\\ 0 & \text{otherwise} \end{cases}$$

Thus, the autocorrelation function is given by

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \begin{cases} 1 & \text{if} \quad \tau = 0\\ -\frac{\theta_1}{1+\theta_1^2} & \text{if} \quad \tau = \pm 1\\ 0 & \text{otherwise} \end{cases}$$

i.e., remained unchanged. Hence, we cannot identify the MA(1) process uniquely from the autocorrelation.

Exercise 12.17 (Moments of AR(1) Process). *Consider a first order AR(1) process without drift:*

$$X_t = \alpha X_{t-1} + \varepsilon_t, \quad |\alpha| < 1$$

(a) Find the mean and the variance
(b) Show that for the variance to be finite, $|\alpha|$ must be less than 1.
(c) Find the autocorrelation function assuming that the process is stationary.

$$\begin{aligned}
X_t &= \alpha X_{t-1} + \varepsilon_t \\
&= \alpha(\alpha X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \alpha^2 X_{t-2} + \alpha\varepsilon_{t-1} + \varepsilon_t \\
&= \alpha^3 X_{t-3} + \alpha^2\varepsilon_{t-2} + \alpha\varepsilon_{t-1} + \varepsilon_t \\
&= \dots \\
&= \sum_{k=0}^{n-1} \alpha^k \varepsilon_{t-k} + \alpha^n X_{t-n}
\end{aligned}$$

Since we have assumed stationarity, i.e., $|\alpha| < 1$, we have: $\alpha^n \longrightarrow 0$ for $n \longrightarrow \infty$. Hence, we can write

$$X_t = \sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k}.$$

(a) It follows that:

$$\begin{aligned}
\mathsf{E}(X_t) &= 0 \\
\mathsf{Var}(X_t) &= \mathsf{Var}\left(\sum_{k=0}^{\infty} \alpha^k \varepsilon_{t-k}\right) = \sum_{k=0}^{\infty} \mathsf{Var}\left(\alpha^k \varepsilon_{t-k}\right) = \sigma^2 \sum_{k=0}^{\infty} \alpha^{2k}
\end{aligned}$$

(b) For the variance to be finite, $|\alpha|$ must be less than 1. In this case we have:

$$\mathsf{Var}(X_t) = \frac{\sigma^2}{1-\alpha^2}.$$

(c) The autocorrelation function is given by

$$\rho(t,\tau) = \frac{\gamma(t,\tau)}{\gamma(t,0)}.$$

$$\begin{aligned}
\gamma(t,\tau) &= \mathsf{Cov}(X_t, X_{t-\tau}) \\
&= \mathsf{E}(X_t X_{t-\tau}) - \mathsf{E}(X_t)\,\mathsf{E}(X_{t-\tau}) \\
&= \mathsf{E}(X_t X_{t-\tau}), \quad \text{since} \quad \mathsf{E}(X_t) = 0 \\
&= \mathsf{E}\{(\alpha X_{t-1} + \varepsilon_t) X_{t-\tau}\} \\
&= \alpha\,\mathsf{E}(X_{t-1} X_{t-\tau}) + \mathsf{E}(\varepsilon_t X_{t-\tau}).
\end{aligned}$$

Since for $\tau > 0$, $X_{t-\tau}$ is a linear combination $\varepsilon_{t-\tau}, \varepsilon_{t-\tau-1}, \ldots$ and therefore uncorrelated with ε_t, we have:

$$\mathsf{E}(\varepsilon_t X_{t-\tau}) = 0$$

and thus, it holds:

$$\gamma_\tau = \alpha\gamma_{\tau-1} = \alpha^2\gamma_{\tau-2} = \ldots = \alpha^\tau\gamma_0.$$

It follows that

$$\rho = \frac{\gamma_\tau}{\gamma_0} = \alpha^\tau.$$

Exercise 12.18 (ARMA(p;q) Representation). *Let X_t be a stationary AR(p) process with mean 0, i.e.*

$$X_t = \sum_{i=1}^{p} \alpha_i X_{t-i} + \varepsilon_t$$

Show that the process Y_t defined as

$$Y_t = \sum_{j=0}^{q} \beta_j X_{t-j}$$

where $\beta_0 = 1$ can be written as an ARMA(p, q) process.

The process Y_t can be written as

$$\begin{aligned}
Y_t &= \sum_{j=0}^{q} \beta_j X_{t-j} \\
&= \sum_{j=0}^{q} \beta_j \left(\sum_{i=1}^{p} \alpha_i X_{t-j-i} + \varepsilon_{t-j} \right)
\end{aligned}$$

$$= \sum_{i=1}^{p} \alpha_i \left(\sum_{j=0}^{q} \beta_j X_{t-j-i} \right) + \sum_{j=0}^{q} \beta_j \varepsilon_{t-j}$$

$$= \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=0}^{q} \beta_j \varepsilon_{t-j}$$

which is exactly an ARMA(p, q) representation.

Exercise 12.19 (Discrete Time and Stationarity).

(a) Let $X(t)$ be a stochastic process in $\mathbb{R}^3$ defined by the vectorial Ornstein-Uhlenbeck equation $dX(t) = AX(t)dt + e_k \sigma(t) dW(t)$ with e_k being the k-th vector in $\mathbb{R}^3$ with $k = 1, 2, 3$. Furthermore, $\sigma(t)$ is a real valued square integrable function and A is the 3×3 matrix.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_3 & -\alpha_2 & -\alpha_1 \end{pmatrix}$$

We suppose that $\alpha_k, k = 1, 2, 3$ are constant. Show that by iterating the finite difference approximations of the time dynamics of the $CAR(3)$ process, we can get the time discrete version for $t = 0, 1, 2$. i.e. $X_1(t+3) \approx (3-\alpha_1)X_1(t+2) + (2\alpha_1 - \alpha_2 - 3)X_1(t+1) + (1 + \alpha_2 - \alpha_3 - \alpha_1)X_1(t)$

(b) The stationarity condition for a $CAR(3)$ model says that the eigenvalues of the matrix A need to have negative real parts. Supposing that $\beta_1 = 0.41, \beta_2 = -0.2, \beta_3 = 0.07$ and using the results from the previous question, verify that the stationarity condition holds.

(a) From the vectorial Ornstein-Uhlenbeck process we define:

$$X_1(t+1) - X_1(t) = X_2(t)dt \tag{12.1}$$

$$X_2(t+1) - X_2(t) = X_3(t)dt \tag{12.2}$$

$$X_3(t+1) - X_3(t) = -\alpha_3 X_1(t)dt - \alpha_2 X_2(t)dt - \alpha_1 X_3(t)dt \tag{12.3}$$

$$X_1(t+2) - X_1(t+1) = X_1(t+1)dt \tag{12.4}$$

$$X_2(t+2) - X_2(t+1) = X_3(t+1)dt \tag{12.5}$$

$$X_3(t+2) - X_3(t+1) = -\alpha_3 X_1(t+1)dt - \alpha_2 X_2(t+1)dt - \alpha_1 X_3(t+1)dt \tag{12.6}$$

$$X_1(t+3) - X_1(t+2) = X_1(t+2)dt \tag{12.7}$$

$$X_2(t+3) - X_2(t+2) = X_3(t+2)dt \tag{12.8}$$

$$X_3(t+3) - X_3(t+2) = -\alpha_3 X_1(t+2)dt - \alpha_2 X_2(t+2)dt \quad (12.9)$$
$$-\alpha_1 X_3(t+2)dt$$

From (12.7) we know that $X_1(t+3) = X_1(t+2)dt + X_1(t+2)$. Substituting (12.4) in (12.5) and setting $dt = 1$, we obtain:

$$\begin{aligned} X_2(t+2) &= X_2(t+1) + X_3(t+1) \\ &= X_2(t+1) + (-\alpha_3 X_1(t) + \alpha_2 X_2(t) \\ &\quad -\alpha_1 X_3(t) + X_3(t)) \end{aligned} \quad (12.10)$$

and from

$$X_3(t) = X_2(t+1) - X_2(t) \quad (12.11)$$
$$X_2(t) = X_1(t+1) - X_1(t) \quad (12.12)$$
$$X_2(t+1) = X_1(t+2) - X_1(t+1) \quad (12.13)$$

Substituting Eqs. (12.11)–(12.13) into (12.10) we have

$$\begin{aligned} X_3(t) &= X_1(t+2) + 2X_1(t+1) + X_1(t) \\ &= X_2(t+2) - X_1(t+2)(2-\alpha_1) + X_1(t+1)(-3-\alpha_2+2\alpha_1) \\ &\quad -X_1(t)(-\alpha_3+\alpha_2-\alpha_1+1) \\ X_1(t+3) &\approx X_1(t+2)(3-\alpha_1) + X_1(t+1)(-3-\alpha_2-2\alpha_1) + X_1(t) \\ &\quad (-\alpha_3+\alpha_2-\alpha_1+1) \end{aligned}$$

(b) From the fitted matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_3 & -\alpha_2 & -\alpha_1 \end{pmatrix}$$

We need to solve the system of equations

$$\begin{aligned} \beta_1 &= 3-\alpha_1 \Rightarrow \alpha_1 = 3-\beta_1 \\ \beta_2 &= 2\alpha_1-\alpha_2-3 \Rightarrow \alpha_2 = 2\alpha_1-3-\beta_2 \\ \beta_3 &= \alpha_2-d_j+d_1+1 \Rightarrow \alpha_3 = \alpha_2-\beta_3-\alpha_1+1 \end{aligned}$$

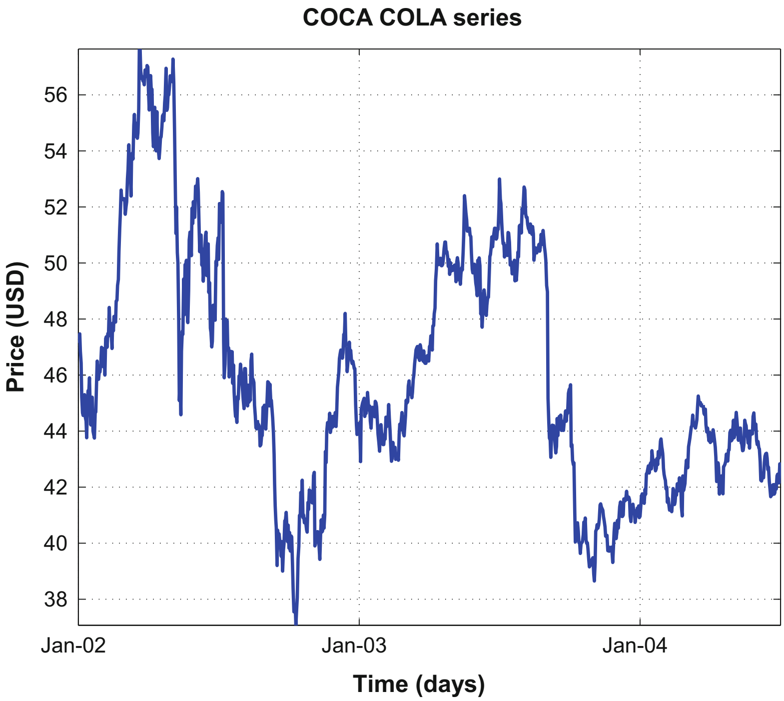


Fig. 12.2 Time plot of the Coca-Cola price series from January 2002 to November 2004 SFScola1

Substituting the values of the α's we obtain $\alpha_1 = 2.09$, $\alpha_2 = 1.38$, $\alpha_3 = 0.22$. The eigenvalues of the fitted matrix therefore are $X_1 = -0.217$, $X_{2,3} = -0.9291 \mp 0.2934$. Thus, the condition for stationarity is fulfilled.

Exercise 12.20 (Applied Time Series Analysis and GARCH). *Consider the data, $COCACOLA.txt$ containing daily prices (p_t) of the Coca-Cola company from January 2002 to November 2004.*

(a) Display the graph of the time series.
(b) Plot the autocorrelation function of the daily price series up to 100 lags and describe the nature of the decay.
(c) Test for stationarity of (p_t) by any suitable procedure.
(d) Plot the rate of returns r_t using, $r_t = \frac{(p_t - p_{t-1})}{p_{t-1}}$ and $r_t = \log p_t - \log p_{t-1}$. Comment on the return pattern.
(e) Model the return rate r_t as a GARCH(1,1) process.

(a) The plot indicates a non-stationary process, or a random walk, characterized by changing stochastic trend and increasing variance (Fig. 12.2).
(b) Time plot of the Coca-Cola series as follows (Fig. 12.3):

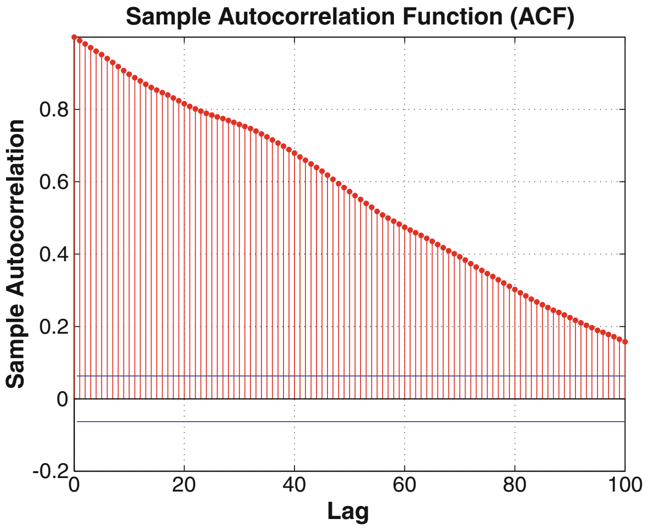


Fig. 12.3 Time plot of Coca-Cola series from January 2002 to November 2004 SFScola2

Table 12.1 Augmented DF test (ADF) for unit root. Critical values are 1 %(−3.458), 5 %(−2.871),10 %(−2.594), see MacKinnon (1991)

Series	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6
ADF test	−2.216	−2.163	−2.304	−2.214	−2.161	−3.458

(c) Testing for unit root is the first step in examining the stationarity of a time series (i.e. testing whether the series are integrated of order 0 $I(0)$ or of order 1 $I(1)$). This is a matter of concern for $(ARIMA)$ modeling and for standard inference procedures for regression models. For example the ADF result in Table 12.1 suggest that (p_t) is non stationary.

(d) Both plots are very similar to each other and show certain common general patterns. There are significant clusters of high variability separated by quieter periods. This changing behavior of the variance is typical for $GARCH$ processes (Fig. 12.4).

(e) The $GARCH(p,q)$ model describes a process where the conditional error variance, σ_t^2 of all information available at time is assumed to obey an $ARMA(p,q)$ model:

$$\sigma_t^2 = \alpha_0 + \alpha_1\sigma_{t-1}^2 + \cdots + \alpha_p\sigma_{t-p}^2 + \beta_1\varepsilon_{t-1}^2 + \cdots + \beta_q\varepsilon_{t-q}^2$$

where ε_t is the error process.

The $GARCH(1,1)$ estimation for r_t is given by

$$\sigma_t^2 = 0.04\sigma_{t-1}^2 + 0.95\varepsilon_{t-1}^2 + \eta_t.$$

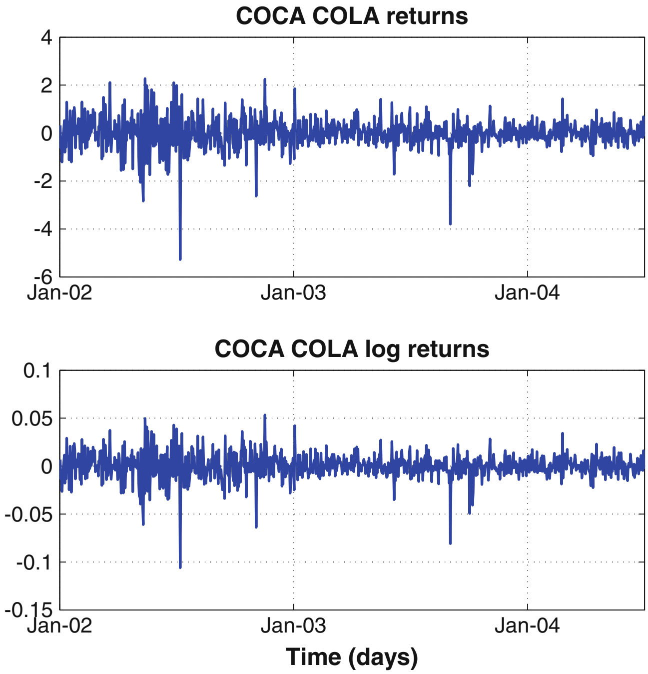


Fig. 12.4 Time plot of Coca-Cola returns from January 2002 to November 2004 `SFScola3`

Exercise 12.21 (GARCH Model). *Given a first order autoregressive model for a series X_t with a $GARCH(1,1)$ process of the error term,*

$$x_t = \theta_0 + \theta_1 x_{t-1} + \varepsilon_t, \tag{12.14}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{12.15}$$

where σ_t^2 is the conditional variance.

(a) Explain how the $GARCH$ model is a generalization of the $ARCH$ model.
(b) Discuss what the model in Eq. (12.15) implies for the process of the squared errors, ε_t^2.
(c) Explain what happens to the model if $\alpha_1 + \beta_1 = 1$ and discuss the implication.

(a) The $GARCH(1,1)$ model is a generalization of the $ARCH(1)$ model by allowing the conditional variance to depend on the lagged conditional variance. Replacing $\sigma_t^2 = \varepsilon_t^2 - \eta_t$ in Eq. (12.15) we obtain:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{12.16}$$

$$\varepsilon_t^2 - \eta_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 (\varepsilon_{t-1}^2 - \eta_{t-1}) \tag{12.17}$$

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + \eta_t - \beta_1\eta_{t-1}, \tag{12.18}$$

This is an $ARMA(1,1)$. Note that the $MA(1)$ part of the squared error process correspond to an infinite autoregressive process, $AR(\infty)$. Hence the $GARCH(1,1)$ is a parsimonious way to model an $ARCH$ process with significant lags, i.e. an $ARCH(p)$ for large p.

(b) The squared error terms, Eq. (12.18),

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + \eta_t - \beta_1\eta_{t-1}$$

follow an $ARMA(1,1)$ where the autoregressive polynomial is given as

$$\theta(L) = 1 - (\alpha_1 + \beta_1)L. \tag{12.19}$$

(c) Following from Eq. (12.19), the persistence of the process depends on the sum $\alpha_1 + \beta_1$. If $\alpha_1 + \beta_1 < 1$, shocks to ε_t^2 have a decaying impact on future volatility. For $\alpha_1 + \beta_1 = 1$, the process has unit roots. Shocks to ε_t^2 will have permanent effect. This model is referred to as, integrated $GARCH$ ($IGARCH$).

Chapter 13
Time Series with Stochastic Volatility

隨機方差時間序列

偷雞不著蝕把米

Try to steal a chicken, but end up with losing the rice

We have already discussed that volatility plays an important role in modeling financial systems and time series. Unlike the term structure, volatility is unobservable and thus must be estimated from market data.

Reliable estimations and forecasts of volatility are important for large credit institutes where volatility is directly used to measure risk. The risk premium, for example, is often specified as a function of volatility. It is interesting to find an appropriate model for volatility. The capability of macroeconomic factors to forecast volatility has already been examined in the literature. Although macroeconomic factors have some forecasting capabilities, the most important factor seems to be the lagged endogenous return. As a result recent studies are mainly concentrated on time series models.

Stock, exchange rates, interest rates and other financial time series have *stylized facts* that are different from other time series. A good candidate for modeling financial time series should represent the properties of stochastic processes. Neither the classic linear AR or ARMA processes nor the nonlinear generalizations can fulfil this task. In this chapter we will describe the most popular volatility class of models: the ARCH (*autoregressive conditional heteroscedasticity*) model that can replicate these stylized facts appropriately.

Exercise 13.1 (Correlation Function). *For the time series of daily DAX and FTSE 100 returns from 1 January 1998 to 31 December 2007, graphically illustrate the following correlation functions:*

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_13,

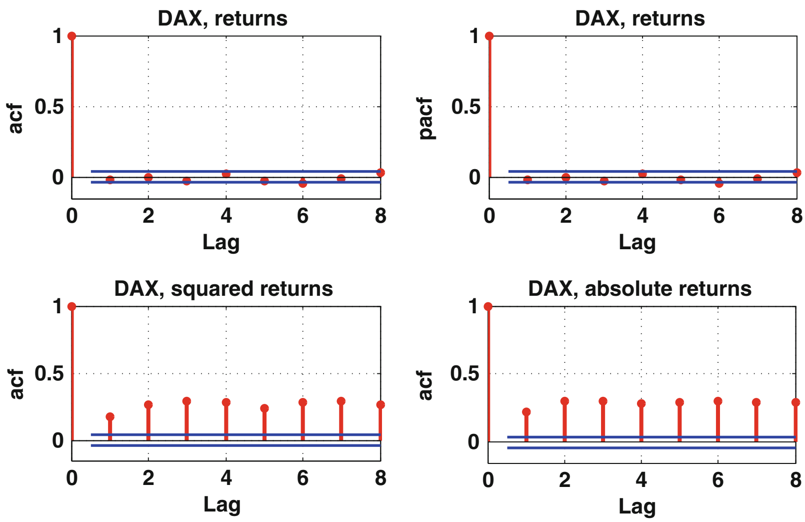


Fig. 13.1 The autocorrelation function and the partial autocorrelation function plots for DAX plain, squared and absolute returns, from 1 January 1998 to 31 December 2007. SFSautoparcorr

1. *Autocorrelation function for plain returns,*
2. *Partial autocorrelation function for plain returns,*
3. *Autocorrelation function for squared returns, and*
4. *Autocorrelation function for absolute returns.*

In addition, compute the Ljung-Box $\left(Q_m^*\right)$ *test statistics, for plain returns, squared returns and absolute returns, as well as the ARCH test statistics for plain returns. Select the number of lags m close to* $\log(n)$, *where n denotes the sample size, see Tsay (2002).*

Are the DAX and FTSE 100 return processes in the period under review:

(a) *Stationary,*
(b) *Serially uncorrelated,*
(c) *Independent?*

Select an appropriate linear time series model for the return processes. Are ARCH and GARCH models appropriate for modeling the volatility processes of the analyzed returns?

The graphical illustration of the empirical autocorrelation functions and the partial correlation functions for analysed time series are given in Figs. 13.1 and 13.2 for the DAX index and the FTSE 100 index, respectively. In the period under review, there are $n = 2,807$ observed returns. By selecting $m = 8$, the computed values for the Ljung-Box Portmanteau statistics and the ARCH test statistics are given in Table 13.1.

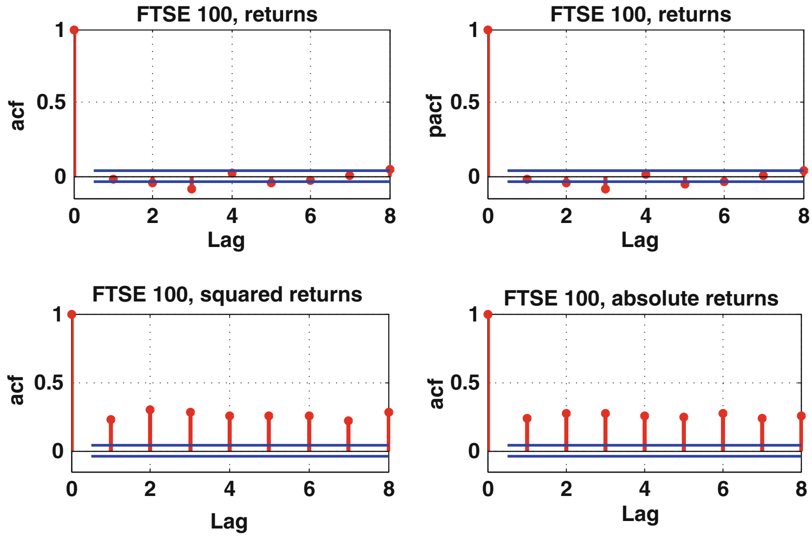


Fig. 13.2 The autocorrelation function and the partial autocorrelation function plots for FTSE 100 plain, squared and absolute returns, from 1 January 1998 to 31 December 2007. SFSautoparcorr

Table 13.1 Ljung-Box (Q_8^*) test statistics (Null hypothesis: $\rho_1 = \ldots = \rho_8 = 0$) for plain, squared and absolute DAX and FTSE 100 returns, as well as the ARCH test statistics (Null hypothesis: no presence of ARCH effects) for plain DAX and FTSE 100 returns, from 1 January 1998 to 31 December 2007. The critical value to reject the null hypothesis is 15.5 at a significance level of 5 % for both tests

	Ljung-Box			ARCH
	r	r^2	$\|r\|$	r
DAX	15.1	1445.0	1728.8	559.6
FTSE 100	41.9	1426.3	1393.7	516.0

The DAX returns are (a) stationary, (b) serially uncorrelated and (c) not independent processes, whereas the FTSE 100 returns are not stationary, but they are serially correlated and dependent processes. An appropriate linear time series model for the DAX returns would be white noise, and for the FTSE 100 returns, for example, an AR(3) model. There is empirical support that the volatility processes are serially correlated. Therefore, it is justified to model the volatility processes of both returns with ARCH or GARCH models.

Exercise 13.2 (Appropriate order q of ARCH(q)). *For modelling of the volatility processes for the DAX and FTSE 100 returns from 1 January 1998 to 31 December 2007, use an ARCH(q) model, $q = 1, \ldots, 15$. The return processes should follow the linear time series models discussed in Exercise 13.1. Based on the value*

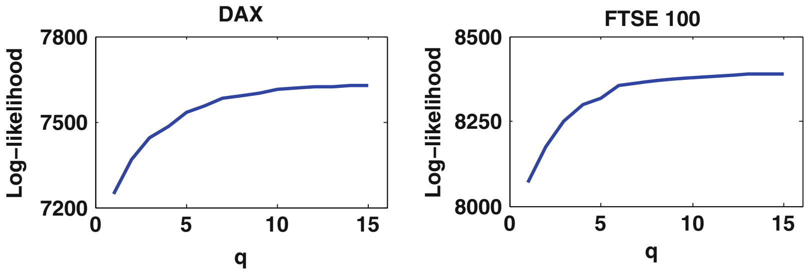


Fig. 13.3 The values of the Log-likelihood function based on the ARCH(q) model for the volatility processes of DAX and FTSE 100 returns, from 1 January 1998 to 31 December 2007. SFSarch

of the optimized log-likelihood objective function select an appropriate order q for modelling and provide the return and volatility equations with estimated parameters.

Furthermore, create a time series of estimated volatility processes in the period under review. Create forecasts of the volatility processes from 1 January 2008 until 31 December 2008 using the unconditional and the conditional volatility approach.

The values of the log-likelihood function based on the ARCH(q) model and computed for different values of q are plotted in Fig. 13.3. Among all fitted ARCH(q) models one observes that an ARCH(6) model is appropriate for modeling both volatility processes in the period under review. Selecting models with higher order does not increase substantially the values of the log-likelihood function.

Return and volatility equations with estimated parameters for the DAX index are:

$$r_t = 7.3 \cdot 10^{-4} + \sigma_t \varepsilon_t \tag{13.1}$$

$$\begin{aligned}\sigma_t^2 = \ & 4.9 \cdot 10^{-5} + 0.045\varepsilon_{t-1}^2 + 0.154\varepsilon_{t-2}^2 + 0.161\varepsilon_{t-3}^2 + 0.126\varepsilon_{t-4}^2 \\ & +0.172\varepsilon_{t-5}^2 + 0.166\varepsilon_{t-6}^2.\end{aligned} \tag{13.2}$$

Return and volatility equations with estimated parameters for the FTSE 100 index are:

$$r_t = 4.2 \cdot 10^{-4} - 0.025 r_{t-1} - 0.039 r_{t-2} - 0.040 r_{t-3} + \sigma_t \varepsilon_t \tag{13.3}$$

$$\begin{aligned}\sigma_t^2 = \ & 2.5 \cdot 10^{-5} + 0.091\varepsilon_{t-1}^2 + 0.130\varepsilon_{t-2}^2 + 0.173\varepsilon_{t-3}^2 + 0.153\varepsilon_{t-4}^2 \\ & +0.126\varepsilon_{t-5}^2 + 0.175\varepsilon_{t-6}^2.\end{aligned} \tag{13.4}$$

The conditional volatility forecast converges asymptotically to the unconditional volatility forecast. As an empirical support of this fact, consider Fig. 13.4.

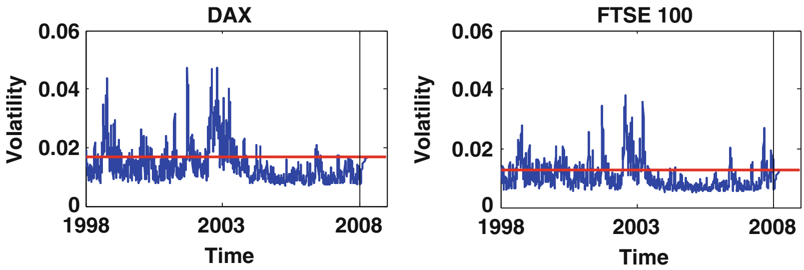


Fig. 13.4 Estimated and forecasted volatility processes of DAX and FTSE 100 returns based on an ARCH(6) model. The *solid line* denotes the unconditional volatility. `SFSarch`

Exercise 13.3 (Specification of the ARCH Model). *Why is a GARCH model sometimes not appropriate to model financial time series? Mention and describe briefly at least one more appropriate specification of the ARCH model.*

There is evidence in the financial markets that a negative shock tends to increase volatility more than a positive shock. Therefore, not only the size of the return but also the sign is important in describing the characteristics of the variance of asset returns.

For example the EGARCH is capable of modeling the described behavior. The volatility of the EGARCH model, which is measured by the conditional variance, is an explicit multiplicative function of lagged innovations. On the contrary, volatility of the standard GARCH model is an additive function of the lagged error terms. Another possible model would be the Threshold ARCH.

Exercise 13.4 (ARCH(1) Process). *Analyze an appropriately parameterized ARCH(1) process. Show that as a model for a financial time series, this process reasonably captures the following stylized facts:*

(a) Heavy tails,
(b) White noise structure,
(c) Volatility clustering.

(a) Let the returns be given by a real-valued stochastic process (ε_t) such that $\mathsf{E}[\varepsilon_t|\mathcal{F}_{t-1}] = 0$ and

$$\mathsf{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2.$$

We impose the conditions $\omega > 0$ and $1 > \alpha > 0$ to ensure that volatility is strictly positive and the return process is stationary.

We show that the returns ε_t have **heavy tails**. By basic calculation using properties of the conditional expectation, we obtain the moments

$$\mathsf{E}[\varepsilon_t^2] = \frac{\omega}{1-\alpha} \text{ and } \mathsf{E}[\varepsilon_t^4] = \frac{3\omega^2}{(1-\alpha)^2}\frac{1-\alpha^2}{1-3\alpha^2}$$

Hence, ε_t has kurtosis

$$\kappa = \frac{\mathsf{E}[\varepsilon_t^4]}{\mathsf{E}[\varepsilon_t^2]^2} = 3\frac{1-\alpha^2}{1-3\alpha^2}.$$

A curve discussion shows that for κ to be positive, α must lie in the interval $(0, 1/3]$. Since $1-\alpha^2 > 1-3\alpha^2 > 0$ for $\alpha \in (0, 1/3]$, we have $\kappa > 3$. Hence, the distribution of ε_t is strictly leptokurtic.

(b) The **white noise** property follows from

$$\mathsf{E}[\varepsilon_t] = \mathsf{E}[\mathsf{E}[\varepsilon_t|\mathcal{F}_{t-1}]] = 0$$

and

$$\mathsf{E}[\varepsilon_t\varepsilon_{t-s}] = \mathsf{E}[\mathsf{E}[\varepsilon_t\varepsilon_{t-s}|\mathcal{F}_{t-1}]] = \mathsf{E}[\mathsf{E}[\varepsilon_t|\mathcal{F}_{t-1}]\varepsilon_{t-s}] = 0, \quad s \geq 1.$$

It is important to note, however, that the returns ε_t are not independent, as the squared returns ε_t^2 are not uncorrelated. Indeed,

$$\begin{aligned}\varepsilon_t^2 &= \sigma_t^2 + \varepsilon_t^2 - \sigma_t^2\\ &= \omega + \alpha\varepsilon_{t-1}^2 + \varepsilon_t^2 - \sigma_t^2,\end{aligned}$$

so that ε_t^2 is an AR(1) process with noise $\varepsilon_t^2 - \sigma_t^2$. We can see this from

$$\begin{aligned}\mathsf{E}[\varepsilon_t^2 - \sigma_t^2] &= \mathsf{E}[\mathsf{E}[\varepsilon_t^2 - \sigma_t^2|\mathcal{F}_{t-1}]]\\ &= \mathsf{E}[\sigma_t^2 - \sigma_t^2]\\ &= 0\end{aligned}$$

and

$$\begin{aligned}\mathsf{E}[(\varepsilon_t^2 - \sigma_t^2)(\varepsilon_{t-s}^2 - \sigma_{t-s}^2)] &= \mathsf{E}[\mathsf{E}[(\varepsilon_t^2 - \sigma_t^2)(\varepsilon_{t-s}^2 - \sigma_{t-s}^2)|\mathcal{F}_{t-1}]]\\ &= \mathsf{E}[\mathsf{E}[(\varepsilon_t^2 - \sigma_t^2)|\mathcal{F}_{t-1}](\varepsilon_{t-s}^2 - \sigma_{t-s}^2)]\\ &= \mathsf{E}[(\sigma_t^2 - \sigma_t^2)(\varepsilon_{t-s}^2 - \sigma_{t-s}^2)]\\ &= 0.\end{aligned}$$

As it was shown in Exercise 12.17, the autocovariance function of ε_t^2 is given by $\gamma(s) = \alpha^s$, $s \geq 1$. It is nonzero by our assumption $\alpha \neq 0$.

(c) The property of **volatility clustering** can be gleaned from the recursive relation

$$\mathsf{Var}(\varepsilon_t|\mathcal{F}_{t-1}) = \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2.$$

If the conditional variance σ_{t-1}^2 of ε_{t-1} has an atypically large realization, then the realization of the conditional variance σ_t^2 of ε_t is also likely to be large. The effect continues *ad infinitum*.

Exercise 13.5 (ARCH(1) Process). *For an ARCH(1) process, show that*

$$\mathsf{E}[\sigma_{t+s}^2|\mathcal{F}_{t-1}] = \frac{1-\alpha^s}{1-\alpha}\omega + \alpha^s\sigma_t^2, \quad s \geq 1.$$

Interpret this result.

We prove the formula by induction on s. For $s = 1$ we have

$$\begin{aligned}\mathsf{E}[\sigma_{t+1}^2|\mathcal{F}_{t-1}] &= \mathsf{E}[\omega + \alpha\varepsilon_t^2|\mathcal{F}_{t-1}] \\ &= \omega + \alpha\sigma_t^2.\end{aligned}$$

Now assume that the formula is true for s. Then

$$\begin{aligned}\mathsf{E}[\sigma_{t+s+1}^2|\mathcal{F}_{t-1}] &= \mathsf{E}[\omega + \alpha\varepsilon_{t+s}^2|\mathcal{F}_{t-1}] \\ &= \omega + \alpha\,\mathsf{E}[\mathsf{E}[\varepsilon_{t+s}^2|\mathcal{F}_{t+s-1}]|\mathcal{F}_{t-1}] \\ &= \omega + \alpha\,\mathsf{E}[\sigma_{t+s}^2|\mathcal{F}_{t-1}] \\ &= \omega + \alpha\left(\frac{1-\alpha^s}{1-\alpha}\omega + \alpha^s\sigma_t^2\right) \\ &= \left(\frac{1-\alpha}{1-\alpha} + \frac{\alpha-\alpha^{s+1}}{1-\alpha}\right)\omega + \alpha^{s+1}\sigma_t^2 \\ &= \frac{1-\alpha^{s+1}}{1-\alpha}\omega + \alpha^{s+1}\sigma_t^2.\end{aligned}$$

This proves the formula.

Recall the interpretation of the conditional expectation $\mathsf{E}[\sigma_{t+s}^2|\mathcal{F}_{t-1}]$ as the best forecast of the future conditional volatility σ_{t+s}^2 given the information at time $t-1$. The above result shows that volatility shocks persist in forecasts of future volatility at the geometric rate α. This persistence is consistent with the model's phenomenon of volatility clustering.

Exercise 13.6 (Representation of a Strong GARCH(p,q) Process). *Assume that $p = q$. Otherwise, we can successively add coefficients $\alpha_{q+1} = 0$ or $\beta_{p+1} = 0$ until the condition is fulfilled. Let $0 < \sum_{i=1}^p(\alpha_i + \beta_i) < 1$. Show that*

$$\sigma_t^2 = \omega\sum_{k=0}^{\infty}\sum_{i_1,\ldots,i_k=1}^{p}(\alpha_{i_1}Z_{t-i_1}^2 + \beta_{i_1})\cdots(\alpha_{i_1}Z_{t-i_1-\ldots-i_k}^2 + \beta_{i_k}), \tag{13.5}$$

where Z_t denotes the innovation ε_t/σ_t. In particular, show that the sum on the right-hand side converges (i.e. is finite) almost surely.

Since $p = q$, we have

$$\begin{aligned}\sigma_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \\ &= \omega + \sum_{i=1}^{p} (\alpha_i Z_{t-i}^2 + \beta_i)\sigma_{t-i}^2 . \end{aligned} \tag{13.6}$$

First, we motivate the formula (13.5). By recursion, we obtain

$$\begin{aligned}\sigma_t^2 &= \omega + \sum_{i=1}^{p} (\alpha_i Z_{t-i}^2 + \beta_i) \left\{ \omega + \sum_{j=1}^{p} (\alpha_i Z_{t-i-j}^2 + \beta_j)\sigma_{t-i-j}^2 \right\} \\ &= \omega \left\{ 1 + \sum_{i=1}^{p} (\alpha_i Z_{t-i}^2 + \beta_i) \right\} + \sum_{i,j=1}^{p} (\alpha_i Z_{t-i}^2 + \beta_i)(\alpha_i Z_{t-i-j}^2 + \beta_j)\sigma_{t-i-j}^2 .\end{aligned}$$

This suggests that σ_t^2 is given by (13.5). In order to work with the infinite (random) sum, we need to check that it converges almost surely. Note that all summands are positive and apply the monotone convergence theorem to obtain

$$\begin{aligned}&\mathsf{E}\left\{ \omega \sum_{k=0}^{\infty} \sum_{i_1,\dots,i_k=1}^{p} (\alpha_{i_1} Z_{t-i_1}^2 + \beta_{i_1}) \cdots (\alpha_{i_1} Z_{t-i_1-\dots-i_k}^2 + \beta_{i_k}) \right\} \\ &\quad = \omega \sum_{k=0}^{\infty} \sum_{i_1,\dots,i_k=1}^{p} \mathsf{E}\left\{ (\alpha_{i_1} Z_{t-i_1}^2 + \beta_{i_1}) \cdots (\alpha_{i_1} Z_{t-i_1-\dots-i_k}^2 + \beta_{i_k}) \right\} \\ &\quad = \omega \sum_{k=0}^{\infty} \sum_{i_1,\dots,i_k=1}^{p} (\alpha_{i_1} + \beta_{i_1}) \cdots (\alpha_{i_1} + \beta_{i_k}) \qquad \text{(i.i.d. assumption)} \\ &\quad = \omega \sum_{k=0}^{\infty} \Bigg\{ \underbrace{\sum_{i=1}^{p} (\alpha_i + \beta_i)}_{0<\cdot<1} \Bigg\}^k \\ &\quad < \infty .\end{aligned}$$

Since the sum is integrable, it must converge almost surely.

By induction on p, we show that (13.5) fulfills the GARCH recursion (13.6). The statement is true for $p = 0$. Now let it be true for $p - 1$. With the abbreviations

$$[\cdots] \stackrel{\text{def}}{=} (\alpha_{i_1} Z_{t-i_1}^2 + \beta_{i_1}) \cdots (\alpha_{i_1} Z_{t-i_1-\dots-i_k}^2 + \beta_{i_k})$$

and

$$[\cdots]_i \stackrel{\text{def}}{=} (\alpha_{i_1} Z^2_{t-i-i_1} + \beta_{i_1}) \cdots (\alpha_{i_1} Z^2_{t-i-i_1-\ldots-i_k} + \beta_{i_k}),$$

we have

$$\omega \sum_{k=0}^{\infty} \sum_{\substack{i_1,\ldots,i_k=1 \\ p \notin \{i_1,\ldots,i_k\}}}^{p} [\cdots] = \omega + \sum_{i=1}^{p-1} (\alpha_i Z_i^2 + \beta_i)\, \omega \sum_{k=0}^{\infty} \sum_{\substack{i_1,\ldots,i_k=1 \\ p \notin \{i_1,\ldots,i_k\}}}^{p} [\cdots]_i . \tag{13.7}$$

Adding to both sides of (13.7) all terms of the form

$$\omega(\alpha_{i_1} Z_{t-i_1} + \beta_{i_1}) \cdots (\alpha_{i_1} Z_{t-i_k-\ldots-i_k} + \beta_{i_k})$$

such that $p \in \{i_1, \ldots, i_k\}$, we obtain (13.6).

Exercise 13.7 (Model Identifiability). *A discussion of a model's identifiability precedes any sound implementation or statistical inference involving the model.*

(a) Using the representation (13.5), specify a GARCH process that does not admit a unique parametrization.
(b) Show that the GARCH(1,1) process

$$\sigma_t^2 = 1 + \frac{1}{4}\varepsilon_{t-1}^2 + \frac{1}{2}\sigma_{t-1}^2$$

and the GARCH(2,2) process

$$\sigma_t^2 = \frac{5}{4} + \frac{1}{4}\varepsilon_{t-1}^2 + \frac{1}{16}\varepsilon_{t-2}^2 + \frac{1}{4}\sigma_{t-1}^2 + \frac{1}{8}\sigma_{t-2}^2$$

are equivalent, i.e. the given relationships are satisfied by the same process $(\sigma_t^2)_{t\in\mathbb{Z}}$. What is noteworthy about the polynomials $p(x) = \alpha_1 x + \alpha_2 x^2$ and $q(x) = 1 - \beta_1 x - \beta_2 x^2$ for the second process? Recast your observation as a hypothesis concerning the identifiability of a general GARCH(p, q) process.

(a) Consider a GARCH(p, q) process $(\sigma_t^2)_{t\in\mathbb{Z}}$ with $\alpha_1 = \cdots = \alpha_p = 0$ and $q > 0$. By the preceding exercise, σ_t^2 has the closed form

$$\begin{aligned} \sigma_t^2 &= \omega \sum_{k=0}^{\infty} \sum_{i_1,\ldots,i_k=1}^{q} \beta_{i_1} \cdots \beta_{i_k} \\ &= \omega \sum_{k=0}^{\infty} \left(\sum_{i=1}^{q} \beta_i \right)^k \\ &= \frac{\omega}{1 - \sum_{i=1}^{q} \beta_i} \end{aligned}$$

and, in particular, is constant. However, σ_t^2 can be parameterized in any of an infinite number of ways; the model is unidentifiable.

(b) By recursion, we obtain

$$\begin{aligned}\sigma_t^2 &= 1 + \frac{1}{4}\varepsilon_{t-1}^2 + \left(\frac{1}{4} + \frac{1}{4}\right)\sigma_{t-1}^2 \\ &= 1 + \frac{1}{4}\varepsilon_{t-1}^2 + \frac{1}{4}\sigma_{t-1}^2 + \frac{1}{4}\left(1 + \frac{1}{4}\varepsilon_{t-2}^2 + \frac{1}{2}\sigma_{t-2}^2\right) \\ &= \frac{5}{4} + \frac{1}{4}\varepsilon_{t-1}^2 + \frac{1}{16}\varepsilon_{t-2}^2 + \frac{1}{4}\sigma_{t-1}^2 + \frac{1}{8}\sigma_{t-2}^2.\end{aligned}$$

Hence, the GARCH(1,1) process satisfies the relation of the GARCH(2,2) process. Since (σ_t^2) is uniquely determined by the parameters ω, $\alpha_1, \ldots, \alpha_q$, $\beta_1, \ldots, \beta_p$, it follows that both processes are equivalent. Generally, a recursion on the terms σ_{t-j}^2 of any GARCH process will produce an equivalent GARCH process of higher order.

We notice that the polynomials $p(x) = \alpha_1 x + \alpha_2 x^2 = \frac{1}{4}x + \frac{1}{16}x^2$ and $q(x) = 1 - \beta_1 x - \beta_2 x^2 = 1 - \frac{1}{4}x - \frac{1}{8}x^2$ share the common root $x = -4$. This suggests that recursion to a higher order produces coefficients $\omega, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p$ such that the polynomials $p(x) = \sum_{i=1}^q \alpha_i x^i$ and $q(x) = 1 - \sum_{j=1}^q \beta_j x^j$ always share a common root. Therefore, we conjecture that a GARCH process is identifiable if and only if the polynomials p and q have no common roots. This turns out to be true. The key to proving the following result is to study the power series expansion of the rational function $p(x)/q(x)$.

Let (ε_t) be a GARCH process with parameters $\omega > 0$ and $\alpha_1, \ldots, \alpha_q$, $\beta_1, \ldots, \beta_p \geq 0$, such that $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$. If the random variables $Z_t^2 = \frac{\varepsilon_t^2}{\sigma_t^2}$ are nondegenerate, if $\alpha_q \neq 0$ or $\beta_p \neq 0$, if $a_i > 0$ for at least one $i \geq 1$, and if the polynomials $p(x) = \sum_{i=1}^q \alpha_i x^i$ and $q(x) = 1 - \sum_{j=1}^q \beta_j x^j$ have no common roots, then (ε_t) is uniquely parametrized.

Exercise 13.8 (Yule-Walker estimator). *GARCH models are typically estimated by a numerical implementation of maximum likelihood methods. This procedure has the disadvantage that it does not yield a closed form estimate and can produce different results depending on the algorithm and its starting value. As an alternative, derive the closed form Yule-Walker moment estimator of the strong GARCH(1,1) process*

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2.$$

(a) Express the process as an ARMA(1,1) process in ε_t^2. Compute the autocorrelations $\rho_{\varepsilon^2}(1)$ and $\rho_{\varepsilon^2}(2)$. Express $\alpha + \beta$ in terms of $\rho_{\varepsilon^2}(1)$ and $\rho_{\varepsilon^2}(2)$.

(b) Rewrite $\rho_{\varepsilon^2}(1)$ and $\rho_{\varepsilon^2}(2)$ as a quadratic equation $\beta^2 - c\beta - 1 = 0$ in β for an appropriate constant c depending on $\alpha + \beta$.

In practice the autocovariances $\rho_{\varepsilon^2}(s)$ can be estimated by

$$\hat{\rho}_{\varepsilon^2}(s) = \frac{\sum_{t=h+1}^{T}\left(\hat{\varepsilon}_t^2 - \hat{\sigma}^2\right)\left(\hat{\varepsilon}_{t-h}^2 - \hat{\sigma}^2\right)}{\sum_{t=1}^{T}\left(\hat{\varepsilon}_t^2 - \hat{\sigma}^2\right)},$$

where the estimated squared residuals $\hat{\varepsilon}_t^2$ and estimated unconditional variance $\hat{\sigma}^2$ are supplied by a preliminary ARMA(1,1) estimation. By plug-in and by the above calculations, we obtain an estimate for $\alpha + \beta$ and therefore for c, β, and α.

(a) Let $v_t = \varepsilon_t^2 - \sigma_t^2$. Then

$$\begin{aligned}
\varepsilon_t^2 &= \sigma_t^2 + v_t \\
&= \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + v_t \\
&= \omega + \alpha\varepsilon_{t-1}^2 + \beta(\varepsilon_{t-1}^2 - v_{t-1}) + v_t \\
&= \omega + (\alpha + \beta)\varepsilon_{t-1}^2 + v_t - \beta(v_{t-1}).
\end{aligned}$$

v_t is a white noise process since

$$\mathsf{E}[v_t] = \mathsf{E}[\mathsf{E}[v_t|\mathcal{F}_{t-1}]] = 0$$

and

$$\mathsf{E}[v_t v_{t-s}] = \mathsf{E}[\mathsf{E}[v_t v_{t-s}|\mathcal{F}_{t-1}]] = \mathsf{E}[\mathsf{E}[v_t|\mathcal{F}_{t-1}]v_{t-s}] = 0, \quad s \geq 1.$$

Hence, ε_t^2 is an ARMA(1,1) process with noise v_t. By the theory of ARMA processes,

$$\rho_{\varepsilon^2}(1) = \frac{(1 - \beta^2 - \alpha\beta)\alpha}{1 - \beta^2 - 2\alpha\beta}$$

and

$$\rho_{\varepsilon^2}(2) = (\alpha + \beta)\rho_{\varepsilon^2}(t, 1).$$

In particular,

$$\alpha + \beta = \frac{\rho_{\varepsilon^2}(2)}{\rho_{\varepsilon^2}(t, 1)}.$$

(b) Rearranging the expressions for $\rho_{\varepsilon^2}(1)$ and $\rho_{\varepsilon^2}(2)$, we get

$$c = \frac{(\alpha + \beta)^2 + 1 - 2\rho_{\varepsilon^2}(1)(\alpha + \beta)}{(\alpha + \beta) - \rho_{\varepsilon^2}(1)}.$$

Exercise 13.9 (Best One-step Forecast). *Consider a GARCH(p, q) process. Compute the best one-step forecast $\widehat{\varepsilon}_{t+1}$ of ε_{t+1} based on $\mathcal{F}_t$. What is the conditional* *variance of the forecast error? Provide a* ***nonconditional*** *confidence interval for* ε_{t+1} *with coverage rate* $1-\alpha$*. Interpret the width of this interval.*

The best forecast of ε_{t+1} given the current information $\mathcal{F}_t$ is the conditional expectation $\widehat{\varepsilon}_{t+1} \stackrel{\text{def}}{=} \mathsf{E}[\varepsilon_{t+1}|\mathcal{F}_t] = 0$. The conditional variance of the forecast error is $\mathsf{Var}(\varepsilon_t - \widehat{\varepsilon}_t|\mathcal{F}_t) = \mathsf{Var}(\varepsilon_t|\mathcal{F}_t) = \sigma_t^2$.

For simplicity, let the innovations be Gaussian, and let $z_{\alpha/2}$ denote the $\alpha/2$-quantile of the standard normal distribution. Then

$$I \stackrel{\text{def}}{=} [-z_{\alpha/2}\sigma_t, z_{\alpha/2}\sigma_t]$$

is a $(1-\alpha)$ confidence interval for the forecast $\widehat{\varepsilon}_{t+1} = 0$. To see this, note that

$$\begin{aligned}
\mathrm{P}[\varepsilon_{t+1} \in I_\alpha] &= \mathsf{E}[\mathbf{1}_{\{\varepsilon_{t+1} \in I_\alpha\}}] \\
&= \mathsf{E}[\mathsf{E}[\mathbf{1}_{\{\varepsilon_{t+1} \in I_\alpha\}}|\mathcal{F}_t]] \\
&= \mathsf{E}[\mathrm{P}[\varepsilon_{t+1} \in I_\alpha|\mathcal{F}_t]] \\
&= \mathsf{E}[1-\alpha] \\
&= 1-\alpha
\end{aligned}$$

by definition of the conditional probability of an event with respect to a σ-algebra. The width of the confidence interval is proportional to σ_t, the conditional standard deviation of the forecast error. The width depends on the most current information and plausibly reflects the volatility clustering exhibited by the model.

Part III
Selected Financial Applications

Chapter 14
Value at Risk and Backtesting

Valor en riesgo y testeo retroactivo
El que busca la verdad corre el riesgo de encontrarla
Anyone who seeks the truth, risks to find it.

Manuel Vicent

Value-at-Risk (VaR) is probably the most commonly known measure for quantifying and controlling the risk of a portfolio. Establishing VaR is of central importance to a credit institute. The description of risk is attained with the help of an "internal model", whose job is to reflect the market risk of portfolios and similar uncertain investments over time. The objective parameter in the model is the probability forecast of portfolio changes over a given period. Whether the model and its technical application correctly identify the essential aspects of the risk, remains to be seen and verified. The backtesting procedure serves to evaluate the quality of the forecast of a risk model by comparing the actual results to those generated with the VaR model. For this the daily VaR estimates are compared to the results from hypothetical trading that are held from the end-of-day position to the end of the next day, the so-called "clean backtesting". The concept of clean backtesting is differentiated from that of "mark-to-market" profit and loss ("dirty $P\&L$") analyses in which intra-day changes are also observed. In judging the quality of the forecast of a risk model it is advisable to concentrate on the clean backtesting.

Exercise 14.1 (Methodologies for Calculating VaR). *Discuss the standard methodologies for calculating VaR and explain how they work. Are there advantages and disadvantages of the presented methods?*

The standard methods are:

Parametric: closed form, or variance/covariance: This methodology estimates VaR using an equation that specifies parameters such as volatility, correlation, delta, and gamma. It is a fast and simple calculation, and extensive historical data are not required; only volatility and a correlation matrix are needed.

DOI 10.1007/978-3-642-33929-5_14, © Springer-Verlag Berlin Heidelberg 2013

The methodology is accurate for linear instruments but less accurate for nonlinear portfolios or for skewed distributions. An example of this approach is given by the Delta-normal model of RiskMetrics (1996).

Monte Carlo: The Monte Carlo methodology estimates VaR by simulating random scenarios and revaluing positions in the portfolio. Extensive historical data are not needed. The method is accurate (if used with a complete pricing algorithm) for all instruments and provides a full distribution of potential portfolio values, not just a specific percentile.

Monte Carlo simulation permits the use of various distributional assumptions (normal, t-distribution, normal mixture, etc.). Thus, it can address the issue of fat tails, or leptokurtosis, but only if market scenarios are generated using appropriate distribution assumptions. A disadvantage of this approach is that it is computationally intensive and time consuming, entailing revaluation of the portfolio under each scenario.

Historical: In the historical methodology, VaR is estimated by taking actual historical rates and revaluing positions for each change in the market. Assuming a complete pricing algorithm is used, the method is accurate for all instruments. The methodology provides a full distribution of potential portfolio values rather than just a specific percentile. The user does not need to make distributional assumptions, although parameter fitting may be performed on the resulting distribution.

Tail risk is incorporated but only if the historical data set includes the tail events. Historical analysis is faster than Monte Carlo simulation because fewer scenarios are used, although it is still somewhat computationally intensive and time consuming. A disadvantage is that a significant daily rate history is required, and sampling far back can create problems if the data are irrelevant to current conditions (for example, if currencies have been devalued). Similarly, scaling far into the future can be difficult. An additional disadvantage is that the results are harder to verify at high confidence levels (99 % and beyond).

Exercise 14.2 (Implementation of VaR). *Name and discuss a few important problems in the implementation of VaR.*

The first problem is the estimation of the parameters of asset return distributions. In real-world applications of VaR, it is necessary to estimate means, variances, and correlations of returns. More generally one needs to specify the joint dependence of the asset returns by parametrizing the joint cdf.

The second problem is the actual calculation of position sizes. A large financial institution may have thousands of loans outstanding. The data base of these loans may not classify them by their riskiness, nor even by their time to maturity, or, a bank may have offsetting positions in foreign currencies at different branches in different locations. A long position in SFR in New York may be offset by a short position in SFR in Geneva; the bank's risk – which we intend to measure by VaR – is based on the net position.

Exercise 14.3 (Expected Shortfall). *Let Z be a $N(0,1)$ rv, prove that* $\vartheta = \mathsf{E}[Z|Z > u] = \varphi(u)/\{1 - \Phi(u)\}$.

Given a threshold u, the exceedances above u are calculated conditional on $\{Z > u\}$ and by using Bayes' rule: $f(z|z > u) = f(z)/\{1 - F(u)\}$, we then obtain:

$$\begin{aligned}
\vartheta = \mathsf{E}[Z|Z > u] &= \int_u^\infty x\varphi(x)\{1-\Phi(u)\}^{-1}dx \\
&= \{1-\Phi(u)\}^{-1}\int_u^\infty x\varphi(x)dx \\
&= \{1-\Phi(u)\}^{-1}\int_u^\infty x(2\pi)^{-1/2}\exp(-x^2/2)dx \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\int_u^\infty x\exp(-x^2/2)dx \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\{-\exp(-x^2/2)\}|_u^\infty \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}[0-\{-\exp(-u^2/2)\}] \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\exp(-u^2/2) = \varphi(u)\{1-\Phi(u)\}^{-1}
\end{aligned}$$

Exercise 14.4 (Expected Shortfall). *Recall the definitions of Z and ϑ for given exceedance level u from Exercise 14.3. Prove that* $\varsigma^2 = \mathsf{Var}[Z|Z > u] = 1 + u\vartheta - \vartheta^2$.

$$\begin{aligned}
\mathsf{Var}[Z|Z > u] &= \mathsf{E}[Z^2|Z > u] - \mathsf{E}^2[Z|Z > u] \\
&= \{1-\Phi(u)\}^{-1}\int_u^\infty x^2(2\pi)^{-1/2}\exp(-x^2/2)dx - \vartheta^2 \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\int_u^\infty x^2\exp(-x^2/2)dx - \vartheta^2 \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\int_u^\infty x\, d\{-\exp(-x^2/2)\} - \vartheta^2
\end{aligned}$$

According to integration by parts, we have:

$$\begin{aligned}
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\left[\{-x\exp(-x^2/2)\}|_u^\infty - \int_u^\infty -\exp(-x^2/2)dx\right] \\
&\quad -\vartheta^2 \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\left\{u\exp(-u^2/2) + \int_u^\infty \exp(-x^2/2)dx\right\} - \vartheta^2 \\
&= \{1-\Phi(u)\}^{-1}(2\pi)^{-1/2}\{u\exp(-u^2/2)\} + \{1-\Phi(u)\}^{-1}\{1-\Phi(u)\} \\
&\quad -\vartheta^2
\end{aligned}$$

$$= u\varphi(u)\{1-\Phi(u)\}^{-1} + \{1-\Phi(u)\}^{-1}\{1-\Phi(u)\} - \vartheta^2$$
$$= 1 + u\vartheta - \vartheta^2$$

Exercise 14.5 (Delta-Normal Model). *Consider a portfolio with 2 stocks. The portfolio has 2,000 EUR invested in S_1 and 4,000 EUR in S_2. Given the following variance-covariance matrix of the returns R_1 and R_2:*

$$\Sigma = \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.06^2 \end{pmatrix}$$

calculate the VaR (at 95 %) for each stock and VaR of the portfolio using the Delta-Normal Model.

Having

$$w = \begin{pmatrix} 2000 \\ 4000 \end{pmatrix}$$

we can compute the product:

$$\sigma^2 = \begin{pmatrix} 2000 & 4000 \end{pmatrix} \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.06^2 \end{pmatrix} \begin{pmatrix} 2000 \\ 4000 \end{pmatrix} = 97,600$$

The standard deviation is $\sigma = \sqrt{97,600} = 312.41$. Since the 95 % quantile of the standard normal distribution is 1.65, the VaR of the portfolio is calculated as

$$VaR = 1.65 \cdot 312.41 = 515.48$$

The VaR for each stock is:

$$VaR_1 = 1.65 \cdot 0.1 \cdot 2000 = 330$$
$$VaR_2 = 1.65 \cdot 0.06 \cdot 4000 = 396$$

We note that the sum is 726 which is greater than the VaR of the portfolio. This is related to the subadditivity issue discussed in Franke et al. (2011).

Exercise 14.6 (Incremental, Marginal and Component VaR). *The partition of the portfolio VaR that indicates how much the portfolio VaR would change approximately if the given component was deleted is called component VaR, or CVaR. It can be calculated with $CVaR_i = w_i \Delta VaR_i$ where ΔVaR_i is the incremental VaR of the position i, i.e. how much the VaR of the portfolio increases if we increase the position i by 1. This value can be calculated directly by revaluating the portfolio or be approximated using the marginal VaR. Calculate both incremental and marginal VaR in the case of Exercise 14.5 (change by 1 for each position) and compare them. Calculate the approximated $CVaR_i$. What do you discover?*

Incremental VaR for S_1:

$$\sigma'^2 = \begin{pmatrix} 2,001 & 4,000 \end{pmatrix} \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.06^2 \end{pmatrix} \begin{pmatrix} 2,001 \\ 4,000 \end{pmatrix} = 97,640$$

$$\sigma' = \sqrt{97,640} = 312.474$$

and hence

$$VaR' = 1.65 \cdot 312.474 = 515.5821$$

the increment is $515.5821 - 515.48 = 0.1021$.

For S_2:

$$\sigma''^2 = \begin{pmatrix} 2000 & 4001 \end{pmatrix} \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.06^2 \end{pmatrix} \begin{pmatrix} 2000 \\ 4001 \end{pmatrix} = 97,629$$

$$\sigma'' = \sqrt{97,629} = 312.4561$$

and hence

$$VaR'' = 1.65 \cdot 312.4561 = 515.5525$$

the increment is $515.5525 - 515.48 = 0.0725$.

The **marginal VaR** is defined as:

$$\frac{\partial VaR}{\partial w_i} = \alpha \frac{\partial \sigma}{\partial w_i}$$

With N uncorrelated stocks we obtain:

$$\frac{\partial \sigma^2}{\partial w_i} = 2 w_i \sigma_i^2 + 2 \sum_{j=1, j \neq i}^{N} w_j \sigma_{ij} = 2 w_i \sigma_i^2.$$

Using $\frac{\partial \sigma^2}{\partial w_i} = 2\sigma \frac{\partial \sigma}{\partial w_i}$ we get:

$$\frac{\partial VaR}{\partial w_{t,i}} = \alpha \frac{\partial \sigma}{\partial w_i} = \alpha \frac{w_i \sigma_i^2}{\sigma}$$

$$\frac{\partial VaR}{\partial w_1} = 1.65 \cdot \frac{2000 \cdot 0.1^2}{312.41} = 0.1056$$

$$\frac{\partial VaR}{\partial w_2} = 1.65 \cdot \frac{4000 \cdot 0.06^2}{312.41} = 0.0761$$

The values are very close to the incremental values, the approximation is good because the change in the position is very small.

$$CVaR_1 = w_1 \cdot \frac{\partial VaR}{\partial w_1} = 211.2608$$

$$CVaR_2 = w_2 \cdot \frac{\partial VaR}{\partial w_2} = 304.2156$$

We discover that $CVaR_1 + CVaR_2 = 515.48 = VaR$.

Exercise 14.7 (Portfolio Management). *Suppose a portfolio manager manages a portfolio which consists of a single asset. The natural logarithm of the portfolio value is normally distributed with an annual mean of 10 % and annual standard deviation of 30 %. The value of the portfolio today is 100 million EUR. Taking VaR as a quantile, answer the following:*

(a) What is the probability of a loss of more than 20 million EUR by year end (i.e., what is the probability that the end-of-year value is less than 80 million EUR)?
(b) With 1 % probability, what is the maximum loss at the end of the year? This is the VaR at 1 %.
(c) Calculate the daily, weekly and monthly VaRs at 1 %.

(a) Denoting the value of the portfolio by ϑ it follows that the logarithm of the portfolio value at time T, ϑ_T, is normally distributed:

$$\log(\vartheta_T) \xrightarrow{\mathcal{L}} \mathrm{N}\left\{\log(\vartheta) + (\mu - \frac{\sigma^2}{2})T, \sigma^2 T\right\}$$

The term $\sigma^2 T/2$ appears due to Itô's Lemma. In our case, $\vartheta = 100$, $\mu = 10\,\%$, $\sigma = 30\,\%$. Thus the end-of-year log of the portfolio value is distributed as

$$\log(\vartheta_T) \xrightarrow{\mathcal{L}} \mathrm{N}(4.66017, 0.3^2)$$

This means that the probability that the end-of-year value of the portfolio is less than 80 is given by the cdf of this distribution. To make the calculation simpler we first transform the above distribution ($\mathrm{N}(4.66017, 0.3^2)$) into the standard normal distribution ($\mathrm{N}(0, 1)$) and obtain the probability of a loss of more than 20 million euro by year end as

$$\Phi(\left[\log(80) - \{\log(100) + 0.1 - 0.3^2/2\}\right]/0.3) = 0.17692.$$

(b) Since the 1 % quantile of the standard normal distribution is -2.32635 (i.e., $\Phi(-2.32635) = 0.01$), we first find the critical portfolio value at the 1 % threshold. Thus,

$$\frac{\log(V) - 4.66017}{0.3} = -2.32635$$

which results in $V = 52.5763$ million EUR as the critical portfolio value at 1 %. Therefore, the maximum loss at the end of the year (annual VaR) at 1 % is $100 - 52.5763 = 47.4237$ million EUR.

(c) The daily, weekly and monthly values for T are 1/250, 5/250 and 21/250, respectively. The corresponding distributions for the daily, weekly and monthly log portfolio values are N(4.60539, 0.00036), N(4.60627, 0.0018) and N(4.60979, 0.00756), respectively. Thus, following the same procedure as in (b), we obtain the following VaR values at 1 %:

$$\begin{aligned} DailyVaR &= 0.2568 \text{ million EUR} \\ WeeklyVaR &= 1.2776 \text{ million EUR} \\ MonthlyVaR &= 5.2572 \text{ million EUR} \end{aligned}$$

Exercise 14.8 (Daily VaR in Delta Normal Framework). *Calculate the daily VaR in a delta normal framework for the following portfolio with the given correlation coefficients. Do the same calculation for the cases of complete diversification and perfect correlation.*

Assets	Estimated daily VaR(EUR)	$\rho_{S,FX}$	$\rho_{B,FX}$	$\rho_{S,B}$
Stocks(S)	400 000.00	−0.10	0.25	0.80
Bonds(B)	300 000.00			
Foreign exchange(FX)	200 000.00			

In the delta normal framework, $VaR = z_\alpha \cdot \sigma$. The daily VaR of the portfolio, in this case with three assets, S, B, and FX, can be therefore calculated by:

$$\begin{aligned} VaR^2 = {} & VaR_S^2 + VaR_B^2 + VaR_{FX}^2 \\ & + 2\rho_{S,FX} VaR_S VaR_{FX} + 2\rho_{B,FX} VaR_B VaR_{FX} + 2\rho_{S,B} VaR_S VaR_B \end{aligned}$$

If $\rho_{S,FX} = -0.10$, $\rho_{B,FX} = 0.25$ and $\rho_{S,B} = 0.80$, then $VaR = 704,273$ EUR.

If complete diversification (i.e., $\rho_{S,FX} = \rho_{B,FX} = \rho_{S,B} = 0$), then $VaR = 538,516$ EUR.

If perfect correlation (i.e., $\rho_{S,FX} = \rho_{B,FX} = \rho_{S,B} = 1$), then $VaR = VaR_S + VaR_B + VaR_{FX} = 900,000$ EUR.

Exercise 14.9 (Daily VaR). *A derivatives portfolio has a current market value of 250 million EUR. Marking this derivatives position, to obtain the market value that would have been obtained on the previous 201 trading days, yields the following worst cases for the daily fall in its value (in million EUR):*

−152	−132	−109	−88	−85	−76	−61	−55	−45	−39
− 37	− 32	− 30	−26	−22	−21	−18	−15	−14	−12

Using the above data, what is the daily VaR on this portfolio at the 1 % threshold? At the 5 % threshold? Comment on the relative accuracy of these two calculations.

Computing the daily VaR on a historical basis is straightforward. 201 trading days, so 200 observations on the fall in value. So the 1 % threshold is the 2nd worst outcome, i.e. −132, the 5 % threshold is the 10th worst outcome, i.e. −39. The 5 % calculation is more accurate than the 1 % calculation. This is because outcomes around the 5 % region are relatively close (−45, −39, −37). The empirical distribution yields enough observations to give a fairly accurate estimate. Around 1 % outcomes are far apart (−152, −132, −109), so in this case the empirical distribution does not yield an accurate estimate.

Exercise 14.10 (Subadditivity of VaR based on Delta-Normal Model). *A risk measure ρ is subadditive when the risk of the total position is less than, or equal to, the sum of the risk of individual portfolios. Intuitively, subadditivity requires that risk measures should consider risk reduction by portfolio diversification effects. Subadditivity can be defined as follows: Let X and Y be random variables denoting the losses of two individual positions. A risk measure ρ is subadditive if the following equation is satisfied.*

$$\rho(X+Y) \le \rho(X) + \rho(Y)$$

Using the above definition, show that the VaR based on the Delta-Normal Model is subadditive.

We know that $VaR = z_\alpha \sigma$ in the Delta-Normal Model. Thus we need to show that $z_\alpha \sigma_{X+Y} \le z_\alpha(\sigma_X + \sigma_Y)$. If two random variables have finite standard deviations, the standard deviations are shown to be subadditive as follows. Let σ_X and σ_Y be standard deviations of random variables X and Y, and let σ_{XY} be the covariance of X and Y. Since $\sigma_{XY} \le \sigma_X \sigma_Y$, the standard deviation σ_{X+Y} of the random variable $X + Y$ satisfies subadditivity as follows.

$$\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_{XY}} \le \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\sigma_X \sigma_Y} = \sigma_X + \sigma_Y$$

$\sigma_{XY} \le \sigma_X \sigma_Y$ can be proved as follows. Let $Z = (Y - \mu_Y) - t(X - \mu_X)$ for a real value t, where μ_X and μ_Y are expectations of X and Y, respectively. Then,

$$\begin{aligned} \mathsf{E}[Z^2] &= t^2\, \mathsf{E}[(X-\mu_X)^2] - 2t\, \mathsf{E}[(X-\mu_X)(Y-\mu_Y)] + \mathsf{E}[(Y-\mu_Y)^2] \\ &= \sigma_X^2 t^2 - 2\sigma_{XY} t + \sigma_Y^2 \end{aligned}$$

Here, let $t = \sigma_{XY}/\sigma_X^2$, then,

$$\mathsf{E}[Z^2] = (\sigma_X^2 \sigma_Y^2 - \sigma_{XY}^2)/\sigma_X^2$$

Since $\mathsf{E}[Z^2] \ge 0$, $\sigma_{XY} \le \sigma_X \sigma_Y$ follows.

Exercise 14.11 (Digital Option). *The digital option (also called the binary option) is the right to earn a fixed amount of payment conditional on whether the underlying asset price goes above (digital call option) or below (digital put option) the strike price. Consider the following two digital options on a stock, with the same exercise* *date T. The first option denoted by A (initial premium u) pays 1,000 if the value of the stock at time T is more than a given U, and nothing otherwise. The second option denoted by B (initial premium l) pays 1,000 if the value of the stock at time T is less than L (with $L < U$), and nothing otherwise. Suppose L and U are chosen such that* $\mathrm{P}(S_T < L) = \mathrm{P}(S_T > U) = 0.008$*, where S_T is the stock price at time T. Consider two traders, trader A and trader B, writing one unit of option A and option B, respectively.*

(a) Calculate the VaR at the 99 % confidence level for each trader.
(b) Calculate the VaR at the 99 % confidence level for the combined position on options A and B.
(c) Is the VaR for this exercise subadditive?

(a) VaR at the 99 % confidence level of trader A is $-u$, because the probability that S_T is more than U is 0.008, which is beyond the confidence level. Similarly, VaR at the 99 % confidence level of trader B is $-l$. This is a clear example of the tail risk. VaR disregards the loss of options A and B, because the probability of the loss is less than one minus the confidence level.
(b) VaR at the 99 % confidence level of this combined position (option A plus option B) is $1,000 - u - l$, because the probability that S_T is more than U or less than L is 0.016, which is more than one minus the confidence level (0.01).
(c) Since the sum of VaR of individual positions (option A and B) is $-u - l$, it is clear that VaR is not subadditive for this exercise.

Exercise 14.12 (Backtesting for correct VaR Model). *In the traffic light approach the backtesting surcharge factor increases with the number of exceptions evaluated on 250 historical returns. Consider a correct 1 % VaR model and assume the independence of the returns. It would mean that the appearance exceptions on 250 days follow a binomial distribution with parameter* $n = 250$ *and* $p = 0.01$*. Calculate the probability that there are more than 4 exceptions.*

Let k be a number of exceptions $k \in \{0, 1, \ldots, 250\}$.

$$P(k > 4) = 1 - P(k \leq 4) = 1 - \sum_{k=0}^{4} \binom{n}{k} p^k (1-p)^{n-k} \approx 0.1019 \qquad (14.1)$$

This example shows that the correct model will yield the backtesting surcharge with a probability of 0.1019.

Exercise 14.13 (Backtesting for incorrect VaR Model). *Similarly to Exercise 14.12 consider an incorrect 1 % VaR model which has the true probability of exception* $p = 0.025$*. Calculate the probability that there are less than 5 exceptions.*

Let k be a number of exceptions $k \in \{0, 1, \ldots, 250\}$.

$$P(k < 5) = \sum_{k=0}^{4} \binom{n}{k} p^k (1-p)^{n-k} \approx 0.25 \tag{14.2}$$

The incorrect model will stay in the green zone with a probability of 0.25.

Exercise 14.14 (Portfolio Management).

(a) *Consider a portfolio which consists of 20 stocks of type A. The price of the stock today is 10 EUR. What is the 95 % 1 year Value-at-Risk (VaR) of this portfolio if the 1 year return arithmetic of the stock R_A is normally distributed $N(0, 0.04)$?*

(b) *Consider again a portfolio of 10 stocks of type A and 20 stocks of type B. The prices of the stocks today are 10 EUR and 5 EUR respectively. The joint distribution of the yearly arithmetic returns follows a 2-dimensional normal distribution $N(\mu, \Sigma)$, where $\mu = (0,0)^T$ and the covariance matrix of the returns R_A and R_B is given by:*

$$\Sigma = \begin{pmatrix} 0.04 & 0.02 \\ 0.02 & 0.08 \end{pmatrix}$$

What is the yearly 95 % VaR of the portfolio in this case?

(a) Let $R_A \sim \mathrm{N}(0, 0.04)$. The value of the portfolio V in 1 year is equal to

$$V = 20 \cdot 10 \cdot R_A$$

We can compute the expected return and variance of the portfolio:

$$\mathsf{E}(V) = \mathsf{E}(20 \cdot 10 \cdot R_A) = 200\,\mathsf{E}(R_A) = 200 \cdot 0 = 0$$

$$\mathsf{Var}(V) = \mathsf{Var}(20 \cdot 10 \cdot R_A) = 200^2\,\mathsf{Var}(R_A) = 200^2 \cdot 0.04 = 1{,}600$$

Since $V \sim \mathrm{N}(0, 1{,}600)$ and the 95 % quantile of the standard normal distribution is $\Phi^{-1}(0.95) = 1.65$, the VaR of the portfolio is calculated as:

$$VaR = 1.65 \cdot \sqrt{1{,}600} = 66$$

(b) Let $R = (R_A, R_B)^\top \sim \mathrm{N}(\mu, \Sigma)$. The value of the portfolio V in 1 year is

$$V = 10 \cdot 10 R_A + 20 \cdot 5 R_B = 100(R_A + R_B)$$

The expected return and variance of the portfolio are:

$$\mathsf{E}(V) = \mathsf{E}\{100(R_A + R_B)\} = 100\,\mathsf{E}(R_A + R_B) = 100 \cdot 0 = 0$$

$$\mathsf{Var}(V) = \mathsf{Var}\{100(R_A + R_B)\} = 100^2(0.04 + 0.08 + 2 \cdot 0.02) = 1{,}600$$

As $V \sim \mathrm{N}(0, 1{,}660)$ and $\Phi^{-1}(0.95) = 1.65$, the VaR of the portfolio is equal:

$$VaR = 1.65 \cdot \sqrt{1,600} = 66$$

Exercise 14.15 (Expectation Derivation). *Let* $X \sim \mathrm{N}(\mu, \sigma^2)$ *be a random variable, c is a constant, please calculate* $\mathsf{E}[\max(X - c, 0)]$.

Let $Z \sim \mathrm{N}(0, 1)$ be a standard normal random variable, then with $\varphi'(z) = -z\varphi(z)$:

$$\begin{aligned}
\mathsf{E}[\max(X-c,0)] &= \int_{-\infty}^{\infty} \max(z\sigma + \mu - c, 0)\varphi(z)dz = \int_{\frac{(c-\mu)}{\sigma}}^{\infty} (\mu - c + \sigma z)\varphi(z)dz \\
&= (\mu - c)\int_{\frac{(c-\mu)}{\sigma}}^{\infty} \varphi(z)dz + \sigma \int_{\frac{(c-\mu)}{\sigma}}^{\infty} z\varphi(z)dz \\
&= (\mu - c)\int_{\frac{(c-\mu)}{\sigma}}^{\infty} \varphi(z)dz - \sigma \int_{\frac{(c-\mu)}{\sigma}}^{\infty} \varphi'(z)dz \\
&= (\mu - c)\left[1 - \Phi\left\{\frac{(c-\mu)}{\sigma}\right\}\right] - \sigma \int_{\frac{(c-\mu)}{\sigma}}^{\infty} \varphi'(z)dz
\end{aligned}$$

Moreover, by $\Phi(z) + \Phi(-z) = 1$ and $\int_{\frac{(c-\mu)}{\sigma}}^{\infty} \varphi'(z)dz = -\varphi\left\{\frac{(c-\mu)}{\sigma}\right\}$, we have:

$$\begin{aligned}
&(\mu - c)\left[\Phi\left\{\frac{(\mu-c)}{\sigma}\right\}\right] + \sigma\varphi\left\{-\frac{(\mu-c)}{\sigma}\right\} \\
&= \sigma\left\{\frac{(\mu-c)}{\sigma}\right\}\Phi\left\{\frac{(\mu-c)}{\sigma}\right\} + \sigma\varphi\left\{-\frac{(\mu-c)}{\sigma}\right\} \qquad (14.3)
\end{aligned}$$

Now put $Y = \frac{(\mu-c)}{\sigma}$, thus (14.3) transforms into

$$\begin{aligned}
&= \sigma Y \Phi(Y) + \sigma\varphi(-Y) \\
&= \sigma\{Y\Phi(Y) + \varphi(-Y)\} \\
&= \sigma\{Y\Phi(Y) + \varphi(Y)\}
\end{aligned}$$

With $\Psi(Y) = Y\Phi(Y) + \varphi(Y)$, we finally obtain:

$$\mathsf{E}[\max(X-c,0)] = \sigma\Psi\left\{\frac{(\mu-c)}{\sigma}\right\}$$

Exercise 14.16 (Expected Shortfall). *Calculate the expected shortfall for* $X \sim \mathrm{N}(\mu, \sigma^2)$.

The expected shortfall $ES_\alpha : \mathsf{E}[X|X > VaR]$, is the expectation of losses given that they exceed the quantile VaR. In terms of the VaR it is given by the conditional expectation

$$\begin{aligned}
\mathsf{E}[X|X > VaR] &= \int_{VaR}^{\infty} x f(x|x > VaR)dx \\
&= \int_{VaR}^{\infty} x \frac{f(x)}{P(X > VaR)} dx \\
&= \int_{VaR}^{\infty} \frac{x f(x)}{1-\alpha} dx \\
&= \int_{VaR=\mu+\sigma z_\alpha} \frac{x \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)dx}{1-\alpha} \qquad (14.4)
\end{aligned}$$

The probability for a loss L lower than the VaR is equal to α, i.e. $\mathrm{P}(L < VaR) = \alpha$. Therefore,

$$\begin{aligned}
\mathrm{P}\left(\frac{L-\mu}{\sigma} < \frac{VaR-\mu}{\sigma}\right) &= \alpha \\
\Phi\left(\frac{VaR-\mu}{\sigma}\right) &= \alpha \\
\frac{VaR-\mu}{\sigma} &= \Phi^{-1}(\alpha) = z_\alpha \\
VaR &= z_\alpha \sigma + \mu
\end{aligned}$$

and using the fact that X is normally distributed the pdf can be written as $f(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$. Now set $t = \frac{x-\mu}{\sigma}$ and rearrange it such that $x = \sigma t + \mu$. After further setting $dt = \frac{dx}{\sigma}$ we substitute into Eq. (14.4) to obtain

$$\begin{aligned}
&= \int_{z_\alpha}^{\infty} \frac{(\sigma t + \mu)\varphi(t)dt}{1-\alpha} && (14.5) \\
&= \frac{\sigma}{1-\alpha} \int_{z_\alpha}^{\infty} t\varphi(t)dt + \frac{\mu}{1-\alpha} \int_{z_\alpha}^{\infty} \varphi(t)dt && (14.6) \\
&= \frac{\sigma}{1-\alpha}(-\varphi(t)|_{z_\alpha}^{\infty}) + \mu && (14.7) \\
&= \frac{\sigma}{1-\alpha}\varphi(z_\alpha) + \mu && (14.8)
\end{aligned}$$

since $\int x \exp(-x^2/2)dx = -\exp(-x^2/2)$ and $\int_{z_\alpha}^{\infty} \exp(t)dt = 1-\alpha$. Therefore, if $X \sim \mathrm{N}(0,1)$ then the expected shortfall

$$\mathsf{E}[X|X > VaR] = \frac{\varphi(z_\alpha)}{1-\Phi(z_\alpha)} \qquad (14.9)$$

Chapter 15
Copulae and Value at Risk

Kopuły i Wartość Narażona na Ryzyko
Chciwy dwa razy traci.
The greedy pay twice.

In order to investigate the risk of a portfolio, the assets subjected to risk (*risk factors*) should be identified and the changes in the portfolio value caused by the risk factors evaluated. Especially relevant for risk management purposes are *negative changes* – the portfolio *losses*. The *Value-at-Risk* (VaR) is a measure that quantifies the riskiness of a portfolio. This measure and its accuracy are of crucial importance in determining the capital requirement for financial institutions. That is one of the reasons why increasing attention has been paid to VaR computing methods.

The losses and the probabilities associated with them (the *distribution of losses*) are necessary to describe the degree of portfolio riskiness. The distribution of losses depends on the joint distribution of risk factors. Copulae are very useful for modelling and estimating multivariate distributions. The flexibilty of copulae basically follows from *Sklar's Theorem*, which says that each joint distribution can be "decomposed" into its marginal distributions and a copula C "responsible" for the dependence structure:

$$F(x_1 \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\}.$$

Exercise 15.1 (Valid Copula Functions). *Are the following functions for $a, b \in [0, 1]$ valid copula functions?*

(a) $C_1(a, b) = a^2/2 + b^2/2 - (a - b)^2/2$
(b) $C_2(a, b) = (|a| + |b| - |a - b|)/2$

(a) It is easy to see that $C_1(a, b) = a^2/2 + b^2/2 - (a - b)^2/2 = a^2/2 + b^2/2 - (a^2 + b^2 - 2ab)/2 = ab$ so one obtains the product copula.
(b) For $a > b$ one has $C_2(a, b) = (a + b - a + b)/2 = b$ and for $a < b$, $C_2(a, b) = a$ which yields the minimum copula, see Franke et al. (2011).

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_15, © Springer-Verlag Berlin Heidelberg 2013

Exercise 15.2 (Calculate Joint Distribution). *Let X_1, X_2 be identically distributed (but not independent) random variables with cdf F. Define the random variables $U_i = 1 - F(X_i)$ for $i = 1,2$ and the joint distribution of $(U_1, U_2)^\top$ be given with copula function C. Calculate the joint distribution of $(X_1, X_2)^\top$ i.e. $\mathrm{P}(X_1 \le s, X_2 \le t)$.*

From the definition:

$$\mathrm{P}(X_1 \le s, X_2 \le t) = \mathrm{P}\{U_1 \ge 1 - F(s), U_2 \ge 1 - F(t)\}.$$

Now using the properties:

$$\begin{aligned}
&\mathrm{P}\{U_1 \ge 1 - F(s), U_2 \ge 1 - F(t)\} + \mathrm{P}\{U_1 \ge 1 - F(s), U_2 < 1 - F(t)\} + \\
&\qquad\qquad\qquad\qquad\qquad\qquad \mathrm{P}\{U_1 < 1 - F(s)\} = 1 \\
&\mathrm{P}\{U_1 \ge 1 - F(s), U_2 < 1 - F(t)\} + \mathrm{P}\{U_1 < 1 - F(s), U_2 < 1 - F(t)\} + \\
&\qquad\qquad\qquad\qquad\qquad\qquad \mathrm{P}\{U_2 \ge 1 - F(t)\} = 1,
\end{aligned}$$

one obtains:

$$\begin{aligned}
\mathrm{P}(X_1 \le s, X_2 \le t) &= 1 - \mathrm{P}\{U_1 \ge 1 - F(s), U_2 < 1 - F(t)\} - \mathrm{P}\{U_1 < 1 - F(s)\} \\
&= 1 - [1 - \mathrm{P}\{U_1 < 1 - F(s), U_2 < 1 - F(t)\} \\
&\quad - \mathrm{P}\{U_2 \ge 1 - F(t)\}] - \mathrm{P}(U_1 < 1 - F(s)) \\
&= C\{1 - F(s), 1 - F(t)\} - \mathrm{P}\{F(t) \ge F(X_2)\} \\
&\quad - \mathrm{P}\{F(s) < F(X_1)\} \\
&= C\{1 - F(s), 1 - F(t)\} + F(s) - F(t) - 1
\end{aligned}$$

Exercise 15.3 (Conditional Distribution Method). *One method of generating random numbers from any copula is the conditional distribution method. Consider a pair of the uniform random variables (U, V) with copula C. Show that*

$$\mathrm{P}(V \le v | U = u) = \frac{\partial}{\partial u} C(u, v).$$

Using the fact that a distribution function is right-continuous we write

$$\begin{aligned}
\mathrm{P}(V \le v | U = u) &= \lim_{\Delta u \to 0} \mathrm{P}(V \le v | u < U \le u + \Delta u) \\
&= \lim_{\Delta u \to 0} \frac{\mathrm{P}(u < U \le u + \Delta u, V \le v)}{\mathrm{P}(u < U \le u + \Delta u)}.
\end{aligned}$$

Since U is a random variable uniformly distributed on the interval $[0,1]$ we have $\mathrm{P}(u < U \le u + \Delta u) = \Delta u$ and we calculate

$$\begin{aligned}
\mathrm{P}(V \le v | U = u) &= \lim_{\Delta u \to 0} \frac{\mathrm{P}(u < U \le u + \Delta u, V \le v)}{\Delta u} \\
&= \lim_{\Delta u \to 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} \\
&= \frac{\partial}{\partial u} C(u, v).
\end{aligned}$$

Exercise 15.4 (Inverse of Conditional Distribution). *Let U and V be two uniform random variables whose joint distribution function is a Clayton copula C. Calculate the inverse of the conditional distribution $V|U$.*

The Clayton copula is given by the formula

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

for $\theta > 0$. Using the copula property from Exercise 15.3 we write

$$\begin{aligned}
\mathrm{P}(V \le v | U = u) &= \frac{\partial}{\partial u} C(u, v) = \frac{\partial}{\partial u} (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \\
&= -\theta u^{-\theta - 1} (-1/\theta) (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta - 1} \\
&= u^{-\theta - 1} (u^{-\theta} + v^{-\theta} - 1)^{-(1+\theta)/\theta} \\
&= (u^{\theta})^{-(1+\theta)/\theta} (u^{-\theta} + v^{-\theta} - 1)^{-(1+\theta)/\theta} \\
&= \{u^{\theta} (u^{-\theta} + v^{-\theta} - 1)\}^{-(1+\theta)/\theta}.
\end{aligned}$$

Solving the equation $q = \frac{\partial}{\partial u} C(u, v)$ for v yields

$$\begin{aligned}
q &= \{u^{\theta} (u^{-\theta} + v^{-\theta} - 1)\}^{-(1+\theta)/\theta} \\
q^{-\theta/(1+\theta)} &= u^{\theta} (u^{-\theta} + v^{-\theta} - 1) \\
u^{-\theta} + v^{-\theta} - 1 &= q^{-\theta/(1+\theta)} u^{-\theta} \\
v^{-\theta} &= q^{-\theta/(1+\theta)} u^{-\theta} - u^{-\theta} + 1 \\
v &= \{(q^{-\theta/(1+\theta)} - 1) u^{-\theta} + 1\}^{-1/\theta}.
\end{aligned}$$

The inverse of the conditional distribution of $(V|U)$ is $\{(q^{-\theta/(1+\theta)} - 1) u^{-\theta} + 1\}^{-1/\theta}$.

Exercise 15.5 (Upper and Lower Tail Dependence). *Calculate coefficients of upper and lower tail dependence for the Gumbel copula (Fig. 15.1).*

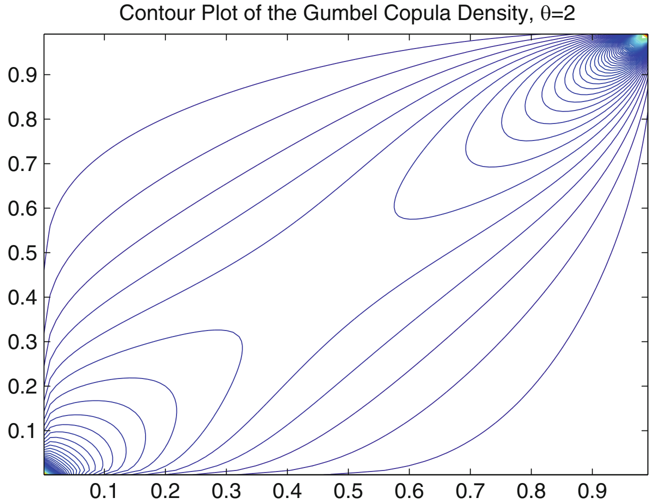


Fig. 15.1 Contour plot of the Gumbel copula density, $\theta = 2$. SFScontourgumbel

The upper and lower tail dependence coefficient is defined as

$$\lambda_U = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1-u} \tag{15.1}$$

$$\lambda_L = \lim_{u \searrow 0} \frac{C(u,u)}{u} \tag{15.2}$$

The Gumbel copula is given by the formula

$$C(u,v) = \exp\left[-\left\{(-\log u)^\theta + (-\log v)^\theta\right\}^{1/\theta} \right].$$

Then

$$\begin{aligned} C(u,u) &= \exp\left[-\left\{2(-\log u)^\theta\right\}^{1/\theta} \right] \\ &= \exp\left(2^{1/\theta} \log u\right) \\ &= \exp\left(\log u^{2^{1/\theta}}\right) \\ &= u^{2^{1/\theta}}. \end{aligned}$$

We now calculate the limits that correspond to the upper and lower tail dependence coefficients for the Gumbel copula. The upper tail dependence coefficient is calculated according to (15.1):

$$\begin{aligned}
\lambda_U &= \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1-u} \\
&= \lim_{u \nearrow 1} \frac{1 - 2u + u^{2^{1/\theta}}}{1-u} \\
&= \lim_{u \nearrow 1} \frac{2 - 2u - 1 + u^{2^{1/\theta}}}{1-u} \\
&= 2 + \lim_{u \nearrow 1} \frac{u^{2^{1/\theta}} - 1}{1-u} \\
&= 2 - \lim_{u \nearrow 1} 2^{1/\theta} u^{2^{1/\theta - 1}} \\
&= 2 - 2^{1/\theta},
\end{aligned}$$

For the lower tail dependence we employ (15.2):

$$\begin{aligned}
\lambda_L &= \lim_{u \searrow 0} \frac{C(u,u)}{u} \\
&= \lim_{u \searrow 0} \frac{u^{2^{1/\theta}}}{u} \\
&= \lim_{u \searrow 0} u^{2^{1/\theta} - 1} \\
&= 0.
\end{aligned}$$

Exercise 15.6 (Copula Application to Finance). *Consider the same situation as in Exercise 14.5, but the asset returns are assumed to be correlated with correlation matrix*

$$R = \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}.$$

(a) Determine the VaR(95 %) *of the portfolio using the Delta-Normal Model and compare the results to the solution of Exercise 14.5.*
(b) Propose a strategy to estimate the VaR *using an elliptical copula with arbitrary margins F_j, $j = 1, 2$.*
(c) Assume the margins follow a student-t distribution with eight degrees of freedom. The simulation based VaR*s are given in Table 15.1. Interpret the results. In which situation should a student-t copula be preferred?*

(a) As the amount I =6,000 EUR is invested with weights $w = (1/3, 2/3)^\top$, the scaled portfolio variance is given by

Table 15.1 Table presents the estimated VaR for different significance levels and computational methods

	Delta-Normal Model	Gaussian copula	Student-t copula
VaR(95 %)	573.58	655.69	674.39
VaR(99 %)	811.23	922.41	1,031.52

$$
\begin{aligned}
1/I^2\sigma^2 &= w^\top \begin{pmatrix} 0.1^2 & \cdot \\ 0.25 * 0.1 * 0.06 & 0.06^2 \end{pmatrix} w \\
&= 0.003378.
\end{aligned}
$$

Then, the VaR of the portfolio can be computed as

$$
\begin{aligned}
\text{VaR} &= 1.65 I \sqrt{0.003378} \\
&= 575.38,
\end{aligned}
$$

which is higher than in the uncorrelated case of Exercise 14.5. This result is reasoned by incorporating the positive dependence between the assets.

(b) The strategy for the copula based methods can be sketched as follows:

(i) Sample n random vectors of the two dimensional copula with correlation matrix R and let $\{u_{ij}\}_{i=1}^{n}$, $j = 1, 2$, be the generated sample.

(ii) Use the additive structure of the portfolio and compute the loss variables

$$
L_i = w_1 F_1^{-1}(u_{i,1})\sigma_{11} + w_2 F_2^{-1}(u_{i,2})\sigma_{22},
$$

where F_j, $j = 1, 2$, are the predetermined or estimated marginal distributions.

(iii) Estimate VaR(95 %) $= \widehat{F}_L^{-1}(0.95)$, where $\widehat{F}_L$ denotes the edf of the loss-variables L_i.

(c) The student-t distributed margins have more probability mass in the tails than the normal distribution, for which reason the Delta-Normal approach must fail to describe the VaR appropriately. The Gaussian copula does not share this shortcoming and permits the margins to follow an arbitrary distribution. The results of Table 15.1 support this property and show, that the copula based VaR(95 %)s are close. In contrast to the student-t copula and irrespective of the correlation $r \in (-1, 1)$, the Gaussian copula cannot describe tail dependency. As a consequence, the student-t copula leads to a more conservative VaR(99 %) than the Gaussian copula, i.e. the student-t copula should be preferred, if the marginal distributions indicate dependence of the tails.

Note, that the results of Table 15.1 rely on generated random vectors and hence, the previous interpretation maybe not hold for a different seed. Table 15.1 and Fig. 15.2 can be reproduced by Quantlet `SFScopapplfin`.

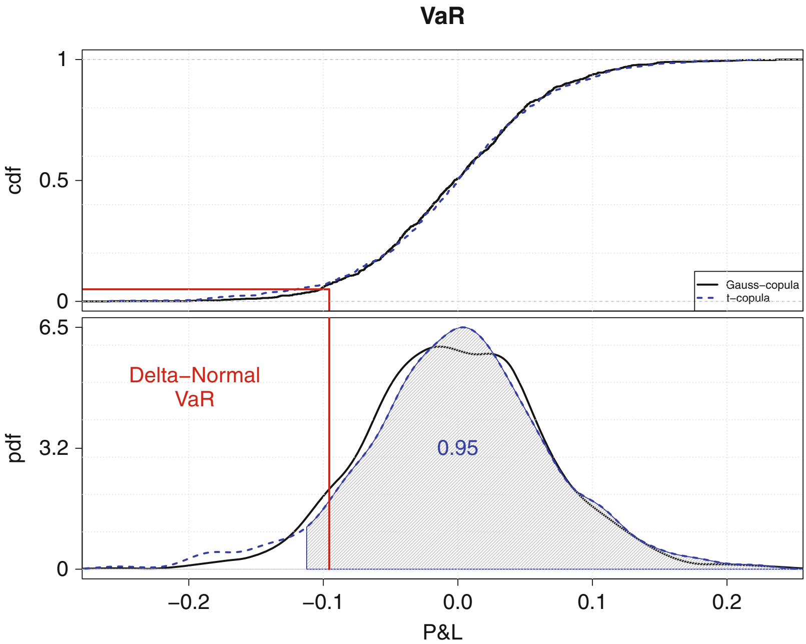


Fig. 15.2 The *upper panel* shows the edfs and the *lower panel* the kernel density estimates of the loss variables for the Gaussian copula (*black solid lines*) and the student-t copula based loss variable (*blue dashed line*). The *red vertical solid line* provides the VaR for the Delta-Normal Model. SFScopapplfin

Chapter 16
Statistics of Extreme Risks

Štatistika extrémnych rizik
Tak dlho sa chodi s džbánom po vodu, kým sa nerozbije
One is going for water with a jug so long until it breaks

When we model returns using a GARCH process with normally distributed innovations, we have already taken into account the second *stylised fact*. The random returns automatically have a leptokurtic distribution and larger losses occur more frequently than under the assumption that the returns are normally distributed. If one is interested in the 95 %-VaR of liquid assets, this approach produces the most useful results. For extreme risk quantiles such as the 99 %-VaR and for riskier types of investments, the risk is often underestimated when the innovations are assumed to be normally distributed, since a higher probability of extreme losses can be produced.

Procedures have therefore been developed which assume that the tails of the innovation's distribution are heavier. Extreme value theory (EVT) plays an important methodological role within the above. The problem we want to solve is how to make statistical inferences about the extreme values in a random process. We want to estimate tails in their far regions and also high quantiles. The key to treating statistics of rare events is the generalised extreme value distribution which also leads to the generalised Pareto distribution. The probability of extreme values will largely depend on how slowly the probability density function $f_Z(x)$ of the innovations goes to 0 as $|x| \to \infty$. However, since extreme observations are rare in data, this produces a difficult estimation problem. Therefore, one has to be supported by extremal event techniques. In this chapter a short overview of the distribution of the extremes and excesses, the return period of some rare events, the frequency of extremal events, the mean excess over a given threshold and several of the latest applications are given.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_16, © Springer-Verlag Berlin Heidelberg 2013

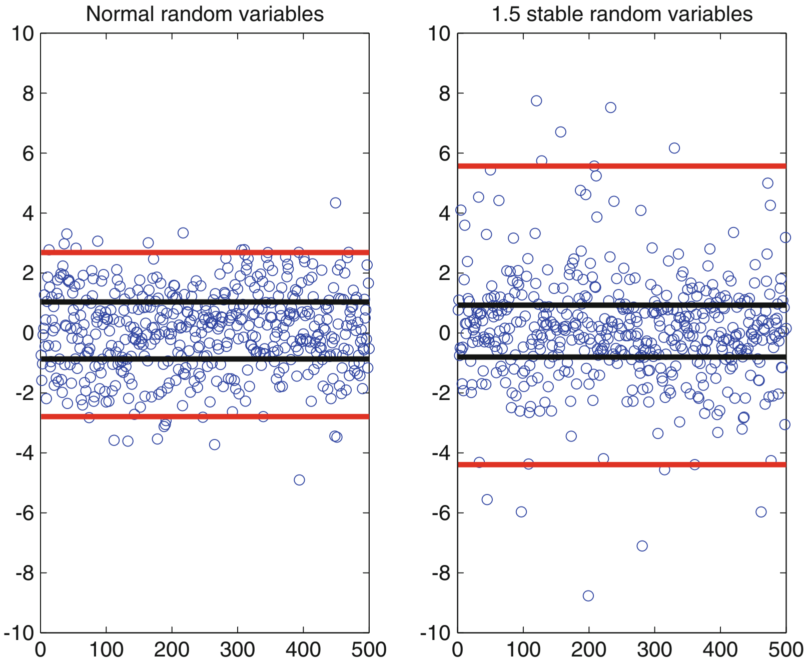


Fig. 16.1 Simulation of 500 1.5-stable and normal variables. `SFSheavytail`

Exercise 16.1 (Tail Behaviour). *The tail behaviour of distributions determines the size and the frequencies of extremal values. Heavy tailed distributions, like stable (including Cauchy and Levy) tend to have more extremal values than distributions like the normal with light exponentially decreasing tails.*

Provide evidence for this statement by simulating stable and normal variates. More precisely, simulate 500 1.5*-stable and normal variables and comment on the size and frequency of the outliers with same scale. For definition and properties of* α*-stable distributions see Cizek et al. (2011, Chap. 1).*

Figure 16.1 displays the simulation result of 500 random normal (left) and 1.5-stable (right) variables. The black lines represent 25 and 75 % quantiles while red lines represent 2.5 and 97.5 % quantiles of the distributions. As we see, the red lines for 1.5-stable random variables are much wider than the ones for normal random variables, which indicates that there are many more extreme values in this case.

Exercise 16.2 (Convergence to Infinity depending on Tail Behaviour). *The maximum of n independent unbounded random variables tends in probability to infinity. The convergence to infinity may be slow or fast depending on the tail behavior of the distributions. Consider a sequence of random variables* $M_n = max(X_1, \ldots, X_n)$, $n = m, 2m, 3m, \ldots$*. Plot n vs.* M_n *for different kinds of distributions. It is suggested*

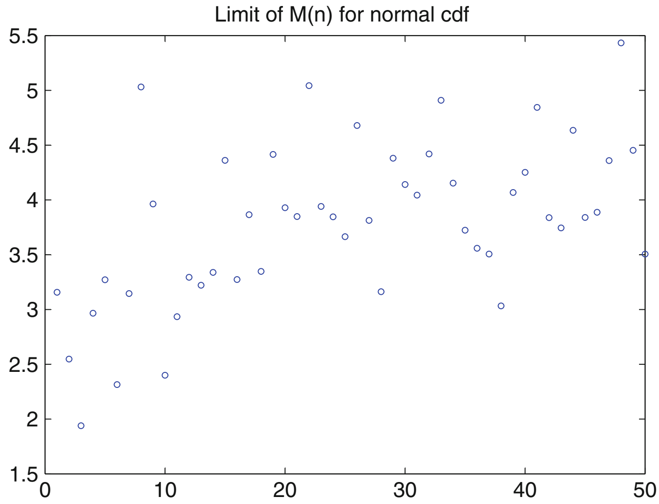


Fig. 16.2 Convergence rate of maximum for n random variables with a standard normal cdf. `SFSmsr1`

that one does this exercise for a standard normal and a stable distribution (see Exercise 16.1).

Let $X_1, \dots, X_n$ are *iid* random variables with a cdf $F(x)$. Then the block maxima $M_n = \max(X_1, \dots, X_n)$ may become arbitrary large. One can easily compute the cdf of maxima:

$$\mathrm{P}(M_n \le x) = \mathrm{P}(X_1 \le x, \dots, X_n \le x) = F^n(x)$$

For unbounded random variables, i.e. $F(x) < 1, \forall x < \infty$,

$$F^n(x) \longrightarrow 0 \text{ and hence } M_n \xrightarrow{\mathrm{P}} \infty$$

First we demonstrate this property for the standard normal distribution, see Fig. 16.2. For the stable distribution the convergence of maximum can be observed on Fig. 16.3.

The rate of convergence to infinity for the (1,1)-stable distributed random variables is higher than for standard normal variables. For $n = 500$, $m = 10$ the difference between two distributions is obvious: the maximum for a standard normal is about 6, and for the stable distributed variables it exceeds 4,500. For the proper analysis of asymptotics one needs the limit law of the maximum domain of attraction (MDA).

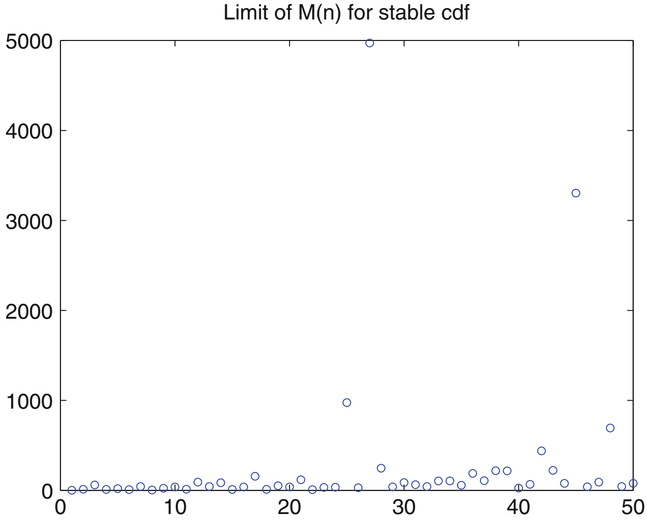


Fig. 16.3 Convergence rate of maximum for n random variables with a 1.1-stable cdf. SFSmsr1

Exercise 16.3 (Asymptotic Distribution). *The empirical quantile is defined on the basis of order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ as $\widehat{x}_q = \widehat{F}_n^{-1}(q)$. Derive the asymptotic distribution of $\widehat{x}_q - x_q$. A good reference is Serfling (2002)*

In Chap. 2 in Serfling (2002) it is shown that $\sqrt{n}(\widehat{x}_q - x_q) \xrightarrow{\mathcal{L}} \mathrm{N}\{0, q(1-q)/f^2(x_q)\}$.

It is not hard to get:

$$f_{X_{(k)}}(x) = n!/\{(k-1)!(n-k)!\}F(x)^{k-1}\{1-F(x)\}^{n-k}f(x).$$

Let $k - 1/n \le 1-q \le k/n$, $\widehat{Z_q} = Z_{(k)}$. When $Z \sim U[0,1]$, $Z_{(k)} \sim B(k, n+1-k)$ (Beta distribution), by the asymptotic property of Beta distribution and CLT:

$$\sqrt{n}(\widehat{Z_q} - q) \xrightarrow{\mathcal{L}} \mathrm{N}\{0, q(1-q)\}.$$

As $X \sim F$, $F(X) \sim U[0,1]$ and $F^{-1}(Z) = X$, by Delta method (Klein, 1977), we have:

$$\sqrt{n}(\widehat{x}_q - x_q) \xrightarrow{\mathcal{L}} \mathrm{N}\{0, q(1-q)/f^2(x_q)\}.$$

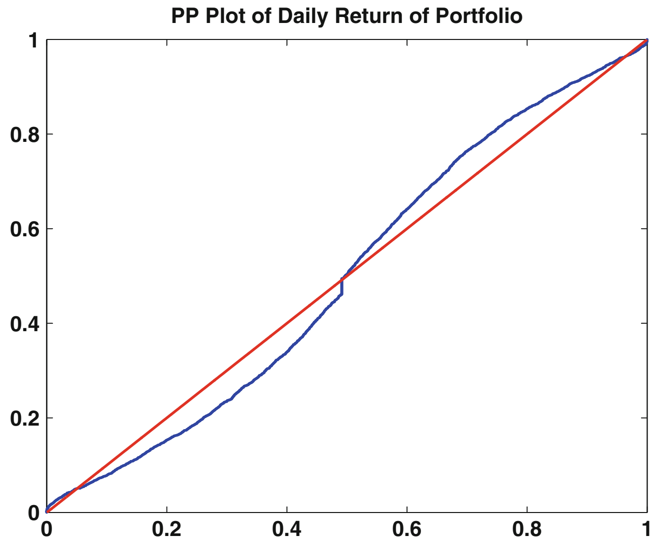


Fig. 16.4 Normal PP plot of daily log-returns of portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29. `SFSportfolio`

Exercise 16.4 (PP-Plot). *The PP-Plot is a diagnostic tool for graphical inspection of the goodness of fit of hypothetical distribution F.*

(a) How is the PP-Plot constructed? Construct the normal PP-Plot of daily log-returns of the portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29 to check the fit of the normal distribution. Is the normal distribution an acceptable approximation of the data?
(b) For the given dataset of the 100 tail values of daily log-returns, estimate the parameter γ using block maxima method. Validate the fit of the GEV distribution with the estimated parameter γ using PP-plot.
(c) Repeat (b) for the Peaks over Threshold (POT) method by estimating γ for the Generalized Pareto Distribution. Use PP-Plot to check the fit of the distribution. Is the approximation better?

(a) The probability-probability plot (PP-plot) is used to checked whether a given data follows some specified distribution. For the normal PP-plot the cumulative probabilities of the data are plotted against the standard normal cdf. It should be approximately linear if the specified distribution is the correct model. For a given dataset of three stocks (Bayer, BMW, Siemens) the corresponding normal PP-Plot is given in Fig. 16.4.
(b) The block maxima method produces a global estimate of $\gamma = 0.0498$. The corresponding PP-plot is depicted in Fig. 16.5.

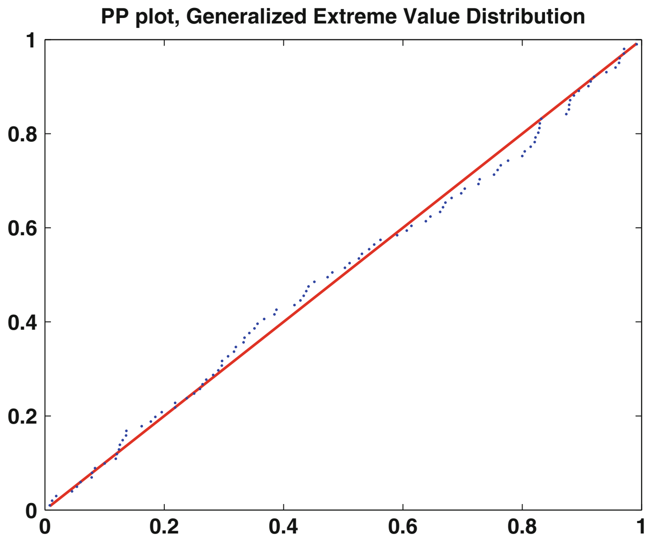


Fig. 16.5 PP plot of 100 tail values of daily log-returns of portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-09-01 against Generalized Extreme Value Distribution with a global parameter $\gamma = 0.0498$ estimated with the block maxima method. SFStailGEV

(c) The POT method gives us a global estimation $\lambda = -0.0768$. Using the PP-Plot in Fig. 16.6 it can be seen that this distribution provides the best approximation of the data.

Exercise 16.5 (QQ-Plot). *The QQ-Plot is a diagnostic tool for graphical inspection of the goodness of fit of hypothetical distribution F.*

(a) What is the advantage of QQ-Plot in comparison with the PP-Plot? Construct the normal QQ-Plot of daily log-returns of the portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29 to check the fit of the normal distribution. Is the normal distribution an acceptable approximation of the data?
(b) For the given dataset of the 100 tail values of daily log-returns globally estimate the parameter γ using block maxima method. Validate the fit of the GEV distribution with the estimated parameter γ using QQ-plot.
(c) Repeat (b) for the Peaks over Threshold (POT) method by estimating a global γ for the Generalized Pareto Distribution. Use QQ-Plot to check the fit of the distribution. Is the approximation better?

(a) QQ-plot is better than the PP plot for assessing the goodness of fit in the tails of the distributions. In order to check how well the normal distribution describes the data, we plot the ordered data $x_{(i)}$ against equally spaced quantiles from a standard normal distribution. For a given dataset of three stocks (Bayer, BMW, Siemens) the results are in Fig. 16.7.

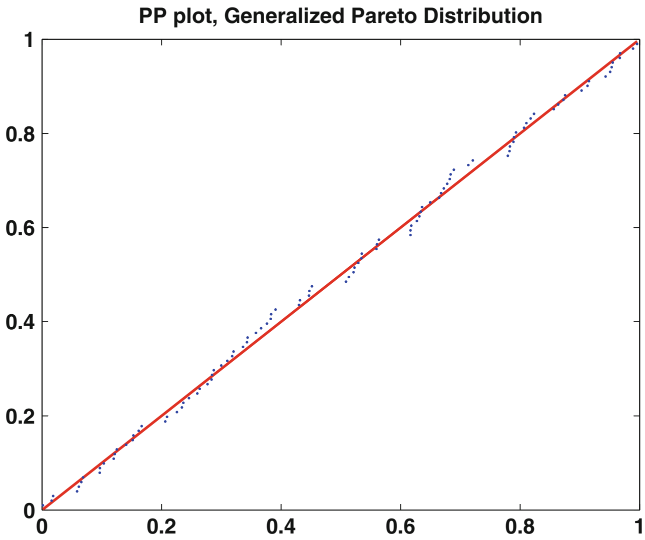


Fig. 16.6 PP plot of 100 tail values of daily log-returns of portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-09-01 against Generalized Pareto Distribution with parameter $\gamma = -0.0768$ globally estimated with POT method. SFStailGPareto

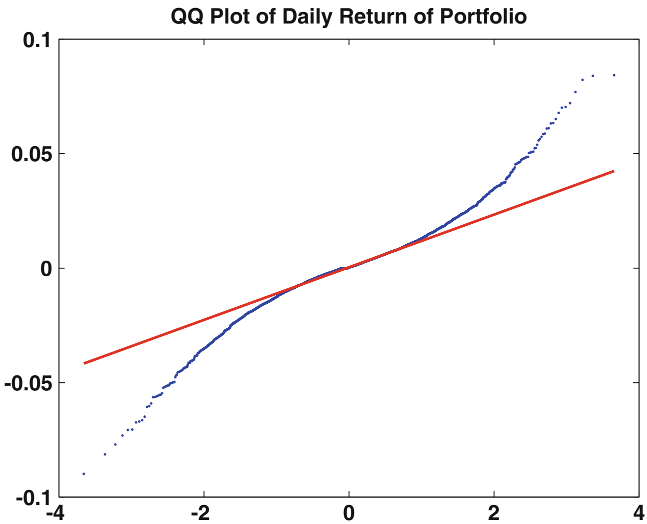


Fig. 16.7 Normal QQ-plot of daily log-returns of portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29. `SFSportfolio`

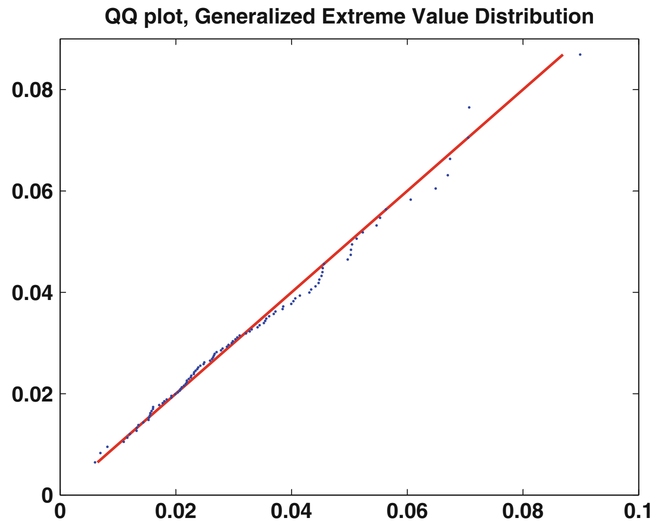


Fig. 16.8 QQ plot of 100 tail values of daily log-returns of portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-09-01 against Generalized Extreme Value Distribution with a global parameter $\gamma = 0.0498$ estimated with the block maxima method. `SFStailGEV`

The closer the line representing the sample distribution is to theoretical distribution, the better is the approximation. As can be seen, the log-returns have much more heavier tails and can not be approximated with the normal distribution.

(b) The block maxima method produces an estimate of a global $\gamma = 0.0498$. The corresponding QQ-plot is depicted in Fig. 16.8:

(c) The POT method gives us a global $\lambda = -0.0768$. Using the QQ-Plot in Fig. 16.9 it can be seen that this distribution provides the best approximation of the data.

Exercise 16.6 (Mean Excess Function). *The mean excess function*

$$e(u) = \mathsf{E}(X - u \mid X > u) \;\; 0 < u < \infty$$

determines not only the tail behavior of the distribution but also uniquely determines F. Prove the formula

$$\overline{F}(x) = \frac{e(0)}{e(x)} \exp\left\{-\int_0^x \frac{1}{e(u)} du\right\}, \quad x > 0$$

Simulate from Fréchet distribution with $\alpha = 2$, do the $PP - plot$ and calculate the mean excess function. Estimate α from the mean excess function and plot the empirical mean excess function.

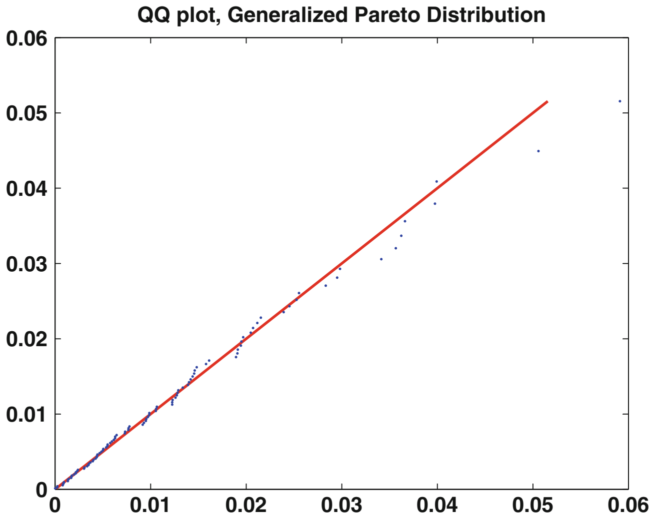


Fig. 16.9 QQ plot of 100 tail values of daily log-returns of portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-09-01 against Generalized Pareto Distribution with a global parameter $\gamma = -0.0768$ estimated with POT method. SFStailGPareto

Suppose that X be a positive, unbounded rv with an absolute continuous cdf F. By definition of $e(u)$, changing the support and using integration by part, we get that

$$\begin{aligned} e(u) &= \mathsf{E}(X-u|X>u) \\ &= \frac{\int_u^\infty x\,dF(x)}{\overline{F}(u)} - u \\ &= \frac{xF(x)|_u^\infty - \int_u^\infty F(x)dx - u\{1-F(x)\}}{\overline{F}(u)} \\ &= \frac{\int_u^\infty 1dx - \int_u^\infty F(x)dx}{\overline{F}(u)} \\ &= \frac{\int_u^\infty \{1-F(x)\}dx}{\overline{F}(u)} \end{aligned}$$

with $\overline{F}(x) = P(X>x) = 1-F(x)$.

We will obtain now an ordinary differential equation, and solve it using the separation of variable method. Steps are shown as follows:

$$\begin{aligned} \overline{F}(u)e(u) &= \int_u^\infty \overline{F}(x)dx \\ \frac{d\{\overline{F}(u)e(u)\}}{du} &= \frac{d\int_u^\infty \overline{F}(x)dx}{du} \end{aligned}$$

$$\begin{aligned}
\overline{F}'(u)e(u) + \overline{F}(u)e'(u) &= -\overline{F}(u) \\
\overline{F}'(u)e(u) &= -\overline{F}(u)\{e'(u) + 1\} \\
\frac{\overline{F}'(u)}{\overline{F}(u)} &= -\frac{e'(u) + 1}{e(u)} \\
\{\log \overline{F}(u)\}' &= -\frac{e'(u) + 1}{e(u)}
\end{aligned}$$

$$\begin{aligned}
\log \overline{F}(u) &= \int_0^u -\frac{e'(x) + 1}{e(x)}\, dx + c \\
&= -\int_0^u \frac{1}{e(x)}\, dx + c - \int_0^u \frac{e'(x)}{e(x)} dx \\
&= -\int_0^u \frac{1}{e(x)}\, dx + c - \log e(x)\Big|_0^u \\
&= -\int_0^u \frac{1}{e(x)}\, dx + c + \log e(0) - \log e(u)
\end{aligned}$$

The general solution is:

$$\begin{aligned}
\overline{F}(u) &= \exp\left\{-\int_0^u \frac{1}{e(x)} dx + c + \log e(0) - \log e(u)\right\} \\
&= \frac{e(0)}{e(u)} \exp\left\{-\int_0^u \frac{1}{e(x)} dx\right\} \cdot c
\end{aligned}$$

Plugging in the boundary condition, we have:

$$\begin{aligned}
\overline{F}(0) &= \frac{e(0)}{e(0)} \exp\left\{-\int_0^0 \frac{1}{e(x)} dx\right\} \cdot c \\
&= 1
\end{aligned}$$

and hence

$$C = 1$$

So we get finally:

$$\overline{F}(u) = \frac{e(0)}{e(u)} \exp\left\{-\int_0^u \frac{1}{e(x)} dx\right\}$$

Therefore a continuous cdf is uniquely determined by its mean excess function. The probability-probability plot (PP-plot) is used to check whether a given data follows some specified distribution. For the normal PP-plot the cumulative probabilities of the data are plotted against the standard normal cdf. Figure 16.10 displays

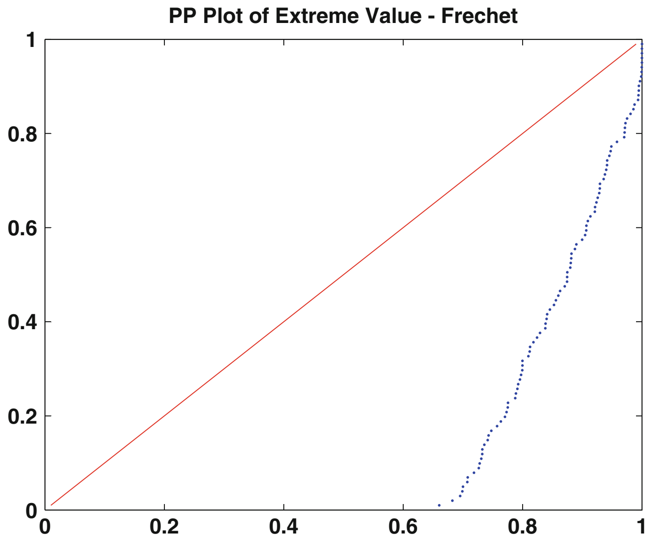


Fig. 16.10 Normal PP plot of the pseudo random variables with Frechét distribution with $\alpha = 2$.
SFSevt2

the Normal PP Plot of the pseudo random variables with Frechét distribution with $\alpha = 2$. The mean excess function of a Fréchet distribution is equal to $e(u) = u\{1 + \mathcal{O}(1)\}/(\alpha - 1)$ and for $u \to \infty$, we observe that $e(u)$ is approximately linear. Figure 16.11 depicts the empirical mean excess function $\widehat{e}_n$, which is estimated based on a representative sample $x_1, \cdots, x_n$: $\widehat{e}_n = \sum_{x_i > x} x_i / \#\{i : x_i > x\}$. In financial risk management, switching from the right tail to the left tail, $e(x)$ is referred to as the expected shortfall. From the empirical mean excess function, ($\alpha = 2$). Observe that Fréchet random variables are away from normal ones (45° line), indicating the presence of heavy tails.

Exercise 16.7 (Pareto Distribution Approximation). *Suppose the Pareto distribution* $\bar{F}(x) = \mathrm{P}(X > x) \sim kx^{-\alpha}$ *where* $\alpha > 0$. *It is well known that an approximation of the parameter* α *can be obtained since* $\log \bar{F}(x) \approx \log k - \alpha \log x$. *Estimate this logarithm approximation for the empirical distribution of the portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29.*

In order to get an approximation $\log \bar{F}(x)$ of the empirical distribution of the portfolio, we first need to estimate the probability $\bar{F}(x)$ for $x = X_{(j)}$ by the relative frequency $\#\{t; X_{t(j)}\}/n = j/n$. Then, we replace $\bar{F}(X_{(j)})$ in $\log \bar{F}(x) \approx \log k - \alpha \log x$ with the estimator j/n. We have then $\log j/n \approx \log k - \alpha \log X_{(j)}$, where $\widehat{\alpha}$ is the slope of the linear regression obtained e.g. by least squares. The linear approximation of $\log \bar{F}(x)$ will only be good in the tails. Thus we estimate $\log \frac{j}{n}$ using only the m biggest order statistics as we observe in Fig. 16.12.

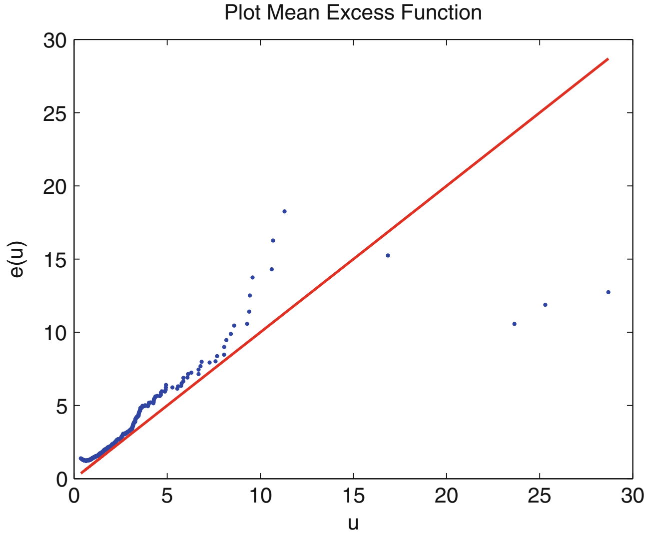


Fig. 16.11 Theoretical (*line*) and empirical (*points*) Mean excess function $e(u)$ of the Frechét distribution with $\alpha = 2$. SFSmef_frechet

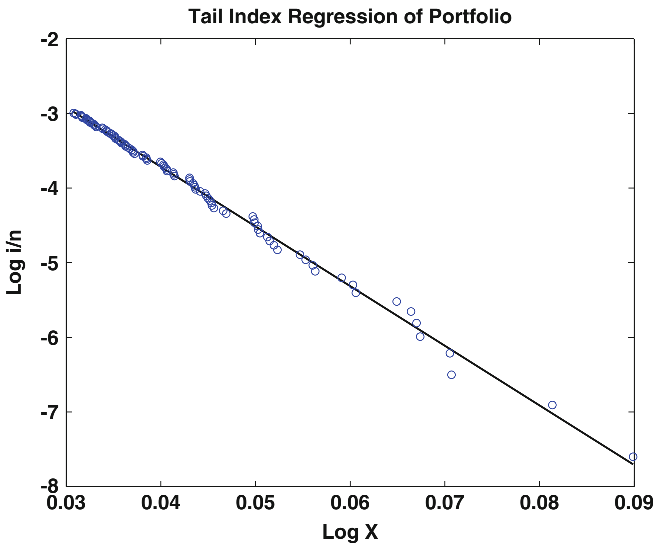


Fig. 16.12 Right tail of the logarithmic empirical distribution of the portfolio (Bayer, BMW, Siemens) negative log-returns from 1992-01-01 to 2006-06-01. SFStailport

Table 16.1 Values of shape Parameter estimated with different methods for the 100 tail observations of the portfolio (Bayer, BMW, Siemens) negative log-returns from 1992-01-01 to 2006-09-01

Method	$\widehat{\gamma}$
Block Max	0.0498
POT	−0.0768
Regression	0.0125
Hill	0.3058

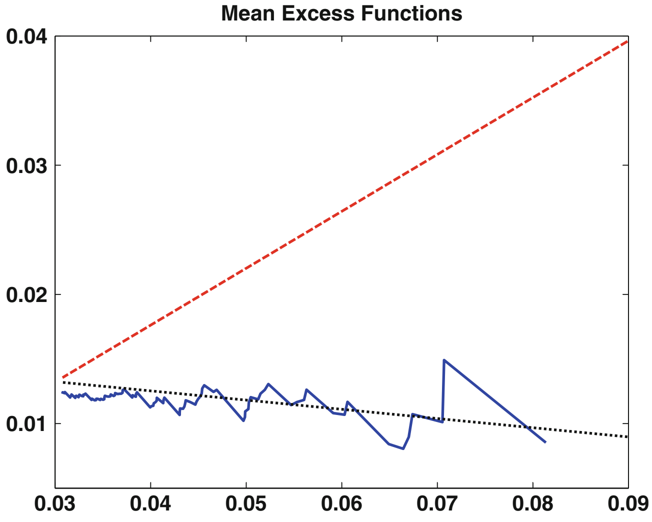


Fig. 16.13 Empirical mean excess plot (*straight line*), mean excess plot of generalized Pareto distribution (*dotted line*) and mean excess plot of Pareto distribution with parameter estimated with Hill estimator (*dashed line*) for portfolio (Bayer, BMW, Siemens) negative log-returns from 1992-01-01 to 2006-09-01. `SFSmeanExcessFun`

Exercise 16.8 (Estimation with Block Max, POT, Hill and Regression). *Estimate the γ parameter locally with Block Maxima, POT, Hill and Regression Model of the portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29. Assuming different types of distributions, plot the mean excess function $e(u)$ for this portfolio. What do you observe?*

The corresponding local shape parameter estimates obtained from the Block Maxima, POT, Hill and Regression Model of the portfolio (Bayer, BMW, Siemens) with 100 observations are given in Table 16.1. The Hill model overstimates α, while the POT underestimates.

We plot in Fig. 16.13, the mean excess plot for the empirical distribution, for the Generalized Pareto distribution, for the Pareto distribution with parameter estimated with Hill estimator for portfolio negative log-returns from 1992-01-01 to 2006-09-01.

Exercise 16.9 (Estimate VaR with Block Max and POT). *Estimate the Value-at-Risk with the Block Maxima Model and with the POT Model of the portfolio (Bayer, BMW, Siemens) from 1992-01-01 to 2006-12-29. Plot the shape and scale parameters estimates over time.*

For a sample of negative returns $\{X_t\}_{t=1}^T$, we decompose the time period 1992-01-01 to 2006-12-29 into k non-overlapping time periods of length 16. We select maximal returns $\{Z_j\}_{j=1}^k$ where $Z_j = \min\{X_{(j-1)n+1}, \dots, X_{jn}\}$ and estimate the parameters of generalized extreme value distribution for the maximal returns $\{Z_j\}_{j=1}^k$. The VaR of the position with given α ($\alpha = 0.95$) in the Block Maxima Model is denoted as:

$$VaR = \mu + \frac{\mu}{\gamma}[\{(1-\alpha^n)\}^{\gamma} - 1]$$

with $\alpha^n = 1 - F(VaR) = \exp\left[-\left\{1 + \gamma\left(\frac{VaR-\mu}{\sigma}\right)^{-1/\gamma}\right\}\right]$.

We use static windows of size $w = 250$ scrolling in time t for VaR estimation $\{X_t\}_{t=s-w+1}^s$ for $s = w, \dots, T$. The VaR estimation procedure generates a time series $\{\widehat{VaR}_{1-\alpha}^t\}_{t=w}^T$ and $\{\hat{\mu}_t\}_{t=w}^T$, $\{\widehat{\sigma}_t\}_{t=w}^T$, $\{\hat{\gamma}_t\}_{t=w}^T$ of parameters estimates. Using Backtesting, one compares the estimated VaR values with true realizations $\{l_t\}$ of the Profit and Loss function to get the ratio of the number of exceedances to the number of observations gives the *exceedances ratio*:

$$\hat{\alpha} = \frac{1}{T-h}\sum_{t=h+1}^{T} \mathbf{1}\{l_t < \widehat{VaR}_{1-\alpha}^t\}$$

Figures 16.14 and 16.15 display the Value-at-Risk estimation under the Block Maxima and the POT Model with 0.05 level for the portfolio formed by Bayer, BMW, Siemens shares during from 1992-01-01 to 2006-09-01. The α-Bactesting result from the Block Maxima model is equal to $\hat{\alpha} = 0.0514$, while for the POT $\hat{\alpha} = 0.0571$. The shape and scale parameter estimates for the Block Maxima model are shown in Fig. 16.16 and for the POT model in Fig. 16.17. In both plots, the threshold of the portfolio is also displayed.

Exercise 16.10 (Calculate VaR). *Let Y_1 be a "short position" in a stock with log-normal distribution*

$$Y_1 = \pi - S$$

with $S = \exp(Z)$ where Z is normally distributed with $N(m, \sigma^2)$.

(a) Calculate $VaR_\alpha(Y_1)$ for $\alpha \in (0,1)$.
(b) Let $Y_1, Y_2, \dots$ independent and identically distributed. Show for $\alpha \in (0,1)$:

$$VaR_\alpha\left(n^{-1}\sum_{i=1}^{n} Y_i\right) \to -\mathsf{E}[Y_1]$$

(c) Which parameter values α violate the convexity property given a large n?

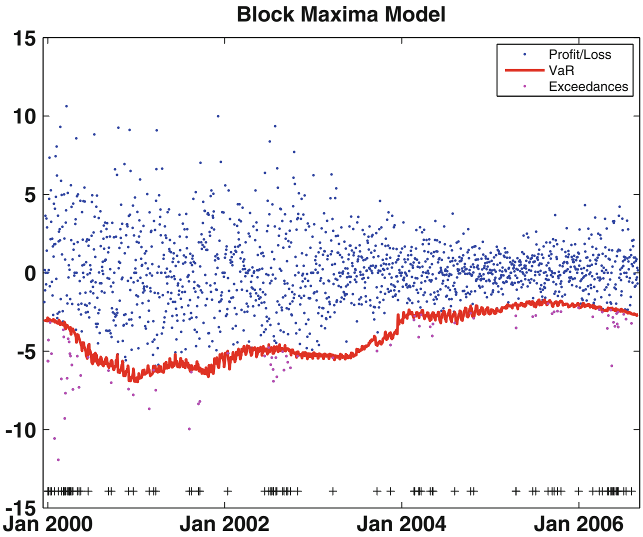


Fig. 16.14 Value-at-Risk estimation at 0.05 level for portfolio: Bayer, BMW, Siemens. Time period: from 1992-01-01 to 2006-09-01. Size of moving window 250, size of block 16. Backtesting result $\hat{\alpha} = 0.0514$. `SFSvar_block_max_backtesting`

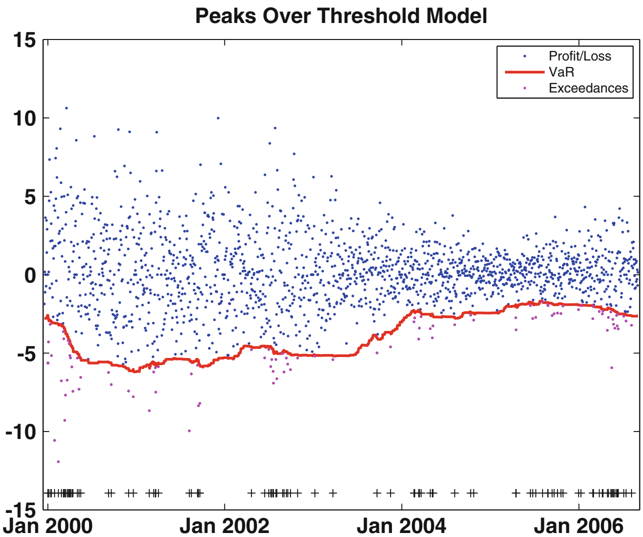


Fig. 16.15 Value-at-Risk estimation at 0.05 level for portfolio: Bayer, BMW, Siemens. Time period: from 1992-01-01 to 2006-09-01. Size of moving window 250. Backtesting result $\hat{\alpha} = 0.0571$. `SFSvar_pot_backtesting`

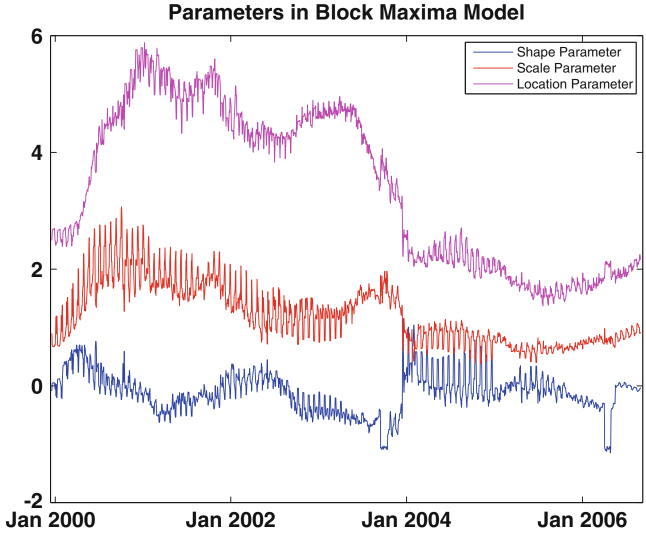


Fig. 16.16 Parameters estimated in Block Maxima Model for portfolio: Bayer, BMW, Siemens. Time period: from 1992-01-01 to 2006-09-01. `SFSvar_block_max_params`

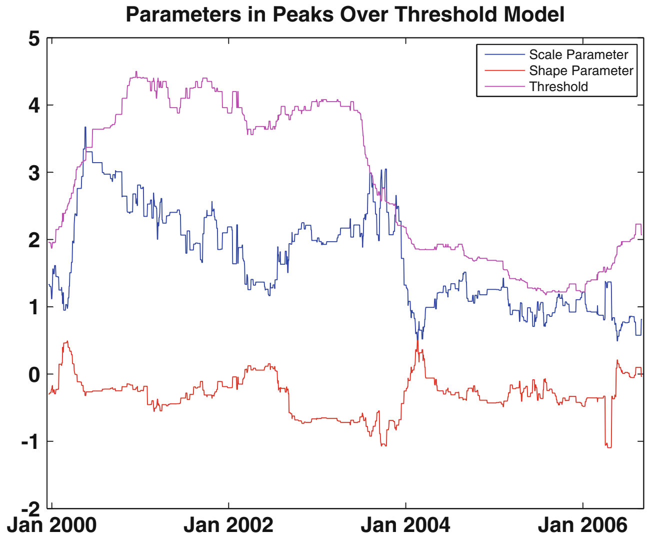


Fig. 16.17 Parameters estimated in POT Model for portfolio: Bayer, BMW, Siemens. Time period: from 1992-01-01 to 2006-09-01. `SFSvar_pot_params`

(a) $VaR_\alpha(Y_1) = \inf\{a|\mathrm{P}(Y_1 + a < 0) \le \alpha\}$

$$\begin{aligned}\mathrm{P}[Y_1 + a < 0] &= \mathrm{P}[\pi - \exp(Z) + a < 0]\\ &= \mathrm{P}[Z > \log(\pi + a)]\\ &= \mathrm{P}\left[(Z - m)/\sigma > \{\log(\pi + a) - m\}/\sigma\right]\end{aligned}$$

Let $X = (Z - m)/\sigma$

$$\begin{aligned}\inf\{a|\mathrm{P}[X > a] \le \alpha\} &= \inf\{a|\mathrm{P}[X \le a] \ge 1 - \alpha\}\\ &= \inf\{a|\Phi(a) \ge 1 - \alpha\}\\ &= \Phi^{-1}(1 - \alpha) = q_{1-\alpha}\end{aligned}$$

Therefore it holds:

$$VaR_\alpha(Y_1) = \inf\{a|(\log(\pi + a) - m)/\sigma = q_{1-\alpha}\} = \exp(\sigma q_{1-\alpha} + m) - \pi$$

(b) Y_i i.i.d. with $\mathsf{E}[Y_1] < \infty$ and $\mathsf{E}[Y_1^2] < \infty$.
Then according to the strong law of large numbers:

$$n^{-1}\sum_{i=1}^{n} Y_i \overset{a.s.}{\rightarrow} \mathsf{E}[Y_1], n \to \infty$$

Hence

$$\mathrm{P}(n^{-1}\textstyle\sum_{i=1}^{n} Y_i < -a) \to \begin{cases} 0 & \mathsf{E}[Y_1] > -a\\ 1 & \mathsf{E}[Y_1] < -a\end{cases}$$

$$VaR_\alpha(n^{-1}\textstyle\sum_{i=1}^{\infty} Y_i) = \inf\left\{a|\mathrm{P}(n^{-1}\sum_{i=1}^{n} Y_i < -a) \le \alpha\right\}$$

$$\inf\{a| - a < \mathsf{E}[Y_1]\} = -\ \mathsf{E}[Y_1]$$

(c) The convexity property is violated if

$$VaR_\alpha(n^{-1}\sum_{i=1}^{n} Y_i) > n^{-1}\sum_{i=1}^{n} VaR_\alpha(Y_i) = VaR(Y_1)$$

$$\mathsf{E}[Y_1] = \pi - \exp(m + \sigma^2/2)$$

Therefore, convexity is not given for big n if

$$-\mathsf{E}[Y_1] = -\pi + \exp(m + \sigma^2/2) > \exp(\sigma q_{1-\alpha} + m) - \pi$$

or equivalently:

$$
\begin{aligned}
\sigma/2 &> q_{1-\alpha} = \Phi^{-1}(1-\alpha) \\
\Phi(\sigma/2) &> 1-\alpha \\
\alpha &> 1-\Phi(\sigma/2) > 0
\end{aligned}
$$

Exercise 16.11 (Normed Risk Measure). *Let ρ be a normed risk measure. Show that*

(a) $\rho(x) \geq -\rho(-x)$ *for all x.*
(b) $\rho(\lambda x) \geq \lambda\rho(x)$ *for all* $\lambda \in (-\infty, 0)$ *and all x.*
(c) For $x \leq 0$ *we have* $\rho(x) \geq 0$.
(d) Let ρ be a normed monetary risk measure. Show that two of the following properties always imply the third property

(α) Convexity
(β) Positive homogeneity
(γ) Subadditivity

Let ρ be a normed and convex risk measure.

(a) We need to show that: $\rho(x) \geq -\rho(-x)\ \forall\ x$

Using the convexity of the risk measure ρ we conclude:

$$
\begin{aligned}
\rho(x)/2 + \rho(-x)/2 &\geq \rho\{(1/2)x + (1/2)(-x)\} \\
&= \rho(0) = 0
\end{aligned}
$$

So we have:

$$\rho(x) \geq -\rho(-x)$$

(b) For $\lambda \in [-1, 0]$ we have:

$$
\begin{aligned}
\rho(\lambda x) &\geq -\rho(-\lambda x) = -\rho\left[-\lambda x + \{1-(-\lambda)\}0\right] \\
&\geq -[\{-\lambda\rho(x) + \{1-(-\lambda)\}\rho(0)] = \lambda\rho(x) \qquad (16.1)
\end{aligned}
$$

For $\lambda \in [-\infty, -1],\ 1/\lambda \in [-1, 0)$, we have:

$$
\begin{aligned}
\rho(1/\lambda \times \lambda x) &\geq \rho(\lambda x)/\lambda \\
\rho(x) &\geq \rho(\lambda x)/\lambda \\
\rho(\lambda x) &\geq \lambda\rho(x)
\end{aligned}
$$

(c) By monoticity of ρ, we have:

$$\rho(x) \geq \rho(0) = 0, \quad \forall\ x \leq 0$$

(d) (i) Let ρ be subadditive and positive homogeneous. This implies convexity. Subadditivity gives:

$$\rho\{\lambda x + (1-\lambda)y\} \leq \rho(\lambda x) + \rho\{(1-\lambda)y\}$$

Positive homogeneity yields:

$$\rho(\lambda x) + \rho\{(1-\lambda)y\} = \lambda\rho(x) + (1-\lambda)\rho(y)$$

Hence ρ is convex $\forall\ \lambda \in [0,1]$ as follows:

$$\rho\{\lambda x + (1-\lambda)y\} \leq \lambda\rho(x) + (1-\lambda)\rho(y)$$

(ii) Let ρ be convex and positive homogeneous. This implies subadditivity.

$$\begin{aligned}(1/2)\rho(x) + (1/2)\rho(y) &\geq \rho\{(1/2)x + (1/2)y\}\\ &= \rho(x+y)/2\end{aligned}$$

Hence ρ is subadditive.

(iii) Let ρ be subadditive and convex. This implies positive homogeneity.
$\lambda = \lceil\lambda\rceil + \tilde{\lambda}$ with $\lceil\lambda\rceil$ the largest integer not larger than λ
$\tilde{\lambda} = \lambda - \lceil\lambda\rceil, \tilde{\lambda} \in [0,1]$

From subadditivity it follows $\rho(nx) \leq n\rho(x)\ \forall\ n \in \mathbb{N}$

So

$$\rho(\lambda x) = \rho(\lceil\lambda\rceil x + \tilde{\lambda}x) \leq \rho(\lceil\lambda\rceil x) + \rho(\tilde{\lambda}x)$$

According to (16.1)

$$\begin{aligned}-\rho(\tilde{\lambda})x &\geq -\tilde{\lambda}\rho(x)\\ \tilde{\lambda}\rho(x) &\geq \rho(\tilde{\lambda}x)\end{aligned}$$

So we have:

$$\rho(\lambda x) = \rho(\lceil\lambda\rceil x + \tilde{\lambda}x) \leq \tilde{\lambda}\rho(x) + \lceil\lambda\rceil\rho(x) = \lambda\rho(x),\ \tilde{\lambda} \in (0,1)$$

$$\rho(\lambda x) \leq \lambda\rho(x) \tag{16.2}$$

On the other hand:

$$\forall\ \lambda \geq 1 : \rho(x) = \rho\{(1/\lambda) \times \lambda x\} \leq \rho(\lambda x)/\lambda \tag{16.3}$$

Hence $\lambda\rho(x) \leq \rho(\lambda x)\ \forall\ \lambda \geq 1$

So $\lambda\rho(x) = \rho(\lambda x)$ by combining (16.2) and (16.3) and
for $\lambda \in (0, 1)$: $\rho(x) = \rho\{(1/\lambda) \times \lambda x\} = \rho(\lambda x)1/\lambda$

Therefore, homogeneity holds for all $\lambda \geq 0$.

Exercise 16.12 (Extreme Losses and EVT). *Suppose that an insurance portfolio has claims X_i which are exponentially distributed* $\exp(\lambda)$. *Suppose that from earlier analysis one has fixed $\lambda = 10$.*

(a) Suppose now that there are $n = 100$ such claims in this portfolio and one observes values larger than 50 and 100. How likely are such extreme losses?
(b) How could you proceed with extreme value theory (EVT)? How could you find the norming constant λ and $\log(n)$*? Does it converge to limit?*

(a) Let $\{X_i\}_{i=1}^n \sim \exp(\lambda), n = 100, \lambda = 10$. The cdf is

$$F(x) = \mathrm{P}(X \leq x) = 1 - \exp(-x/10), x \geq 0$$

Define $M_n = \max(X_1, \ldots, X_n)$. The probability of extreme events for $n = 100$ is calculated as:

$$\begin{aligned}\mathrm{P}(M_{100} > x) &= 1 - \{F(x)\}^{100}\\ &= 1 - \{1 - \exp(-x/10)\}^{100}\end{aligned}$$

If we plug in $x = 50$ and 100 respectively, we obtain 0.4914 and $0.453 \cdot 10^{-2}$ respectively.

(b) Using EVT with an MDA of the Gumbel distribution we find using the correct scaling variables:

$$\begin{aligned}\mathrm{P}\{M_n/10 - \log(n) \leq x\} &= \mathrm{P}[M_n \leq 10\{x + \log(n)\}]\\ &= F^n[10\{x + \log(n)\}]\\ &= \{1 - \exp(-x)/n\}^n\end{aligned}$$

This leads asymptotically to

$$\Lambda(x) = \exp\{-\exp(-x)\}$$

the Gumbel distribution. Using the asymptotic approximation one obtains

$$\mathrm{P}(M_n \leq x) \approx \Lambda[\{x + \log(n)\}\, 10]$$

and therefore $P(M_{100} > 50) \approx 0.4902$, $P(M_{100} > 100) \approx 0.453 \cdot 10^{-2}$.

Exercise 16.13 (Slowly Varying Function). *Show that the function* $L(x) = \log(1 + x)$ *is a slowly varying function.*

Hint: a positive measurable function L in $(0, \infty)$ *that satisfies* $\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1$ *for all* $t > 0$ *is called a slowly varying function.*

Apply Taylor expansion directly to $\log(1 + tx)$ at $t = 1$. We have:

$$\log(1 + tx) = \log(1 + x) + (t - 1)/(1 + x) + \mathcal{O}\{1/(1 + x)^2\}$$
$$\frac{\log(1 + tx)}{\log(1 + x)} = 1 + \mathcal{O}\{\log(1 + x)/(1 + x)\}$$

which is 1 when $x \to \infty$.

Exercise 16.14 (Pareto Distribution). *Let* $X_1, \ldots, X_n$ *are i.i.d. random variables with a Pareto distribution with the cdf*

$$W_{(1,\alpha)}(x) = 1 - \frac{1}{x^\alpha}, x \geq 1, \alpha > 0$$

(a) Calculate $\mathsf{E}\left\{W_{(1,\alpha)}(X)\right\}$.
(b) What is the cdf of $\min(X_1, \ldots, X_n)$*?*

(a) Let $X_1, \ldots, X_n$ have some distribution function F. By definition we know that $F(X) \sim \mathrm{U}(0, 1)$. Thus, $W_{(1,\alpha)}(X)$ follows a uniform distribution $(0, 1)$. The expected value therefore is $\mathsf{E}\left[W_{(1,\alpha)}(X)\right] = \frac{1}{2}$.

(b)

$$\begin{aligned}
\mathrm{P}\{\min(X_1, \ldots, X_n) < t\} &= 1 - \mathrm{P}\{\min(X_1, \ldots, X_n) > t\} \\
&= 1 - \mathrm{P}(X_1 > t, X_2 > t, \ldots, X_n > t) \\
&= 1 - \mathrm{P}(X_1 > t)\mathrm{P}(X_2 > t)\ldots\mathrm{P}(X_n > t) \\
&= 1 - \{\mathrm{P}(X_1 > t)\}^n = 1 - \{1 - F(t)\}^n \\
&= 1 - \left(\frac{1}{X^\alpha}\right)^n
\end{aligned}$$

Exercise 16.15 (Quantile and Expectile Function).

(a) Compare the quantile with the expectile function.
(b) Plot quantile and expectile curves for the normal and uniform distribution.
(c) Construct confidence bands for the expectile functions.

(a) Both quantile and expectile are used to capture the tail behaviours of a distribution, therefore they both are widely applied in financial studies, such as to calculate VaR. Quantile is defined in L_1 norm, while expectile is defined in L_2 norm. Define the contrast function

$$\rho_\tau(u) = |u|^\alpha |\tau - \mathbf{I}(u < 0)|, \qquad 0 < \tau < 1. \tag{16.4}$$

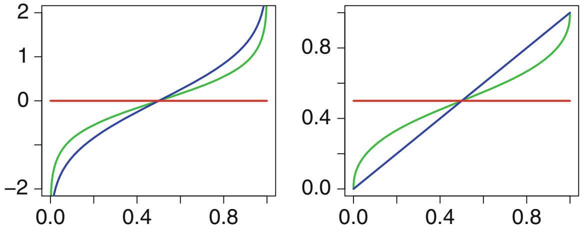


Fig. 16.18 Quantile curve (*blue*) and expectile curve (*green*) for N(0, 1) (*left*) and U(0, 1) (*right*). SFSconfexpectile0.95

with $\alpha = 1, 2$. Minimizing the expectation of the contrast functions, we obtain the τ-th quantile and expectile respectively.

$$l_\tau = \arg\min_\theta \mathsf{E}[\rho_\tau(Y - \theta)], \tag{16.5}$$

for $\alpha = 1$, l_τ gives us the quantile, and l_τ represents the expectile when $\alpha = 2$.

(b) See Fig. 16.18.

(c) We generate bivariate random variables $\{(X_i, Y_i)\}_{i=1}^n$ with sample size $n = 500$. The covariate X is uniformly distributed on $[0, 2]$

$$Y = 1.5X + 2\sin(\pi X) + \varepsilon \tag{16.6}$$

where $\varepsilon \sim \mathrm{N}(0, 1)$.

Obviously, the theoretical expectiles (fixed τ) are determined by

$$v(x) = 1.5x + 2\sin(\pi x) + v_N(\tau) \tag{16.7}$$

where $v_N(\tau)$ is the τ-th expectile of the standard Normal distribution (Figs. 16.19 and 16.20).

Exercise 16.16 (Confidence Band). *Quantile regression is a straightforward way to search for the relation implied in one asset return to other variables. Since the VaR is just the one-step ahead quantile estimation (or prediction) for the distribution of the asset we care, this motivates the use of the quantile regression. However, the functional form is usually unknown, and this increases the difficulty in using the quantile regression. In this exercise, you are asked to construct the confidence band for a nonparametric quantile regression curve, by using the weekly returns of the Bank of America (BOA) and Citigroup (C) from January 31, 2005 to January 31, 2010, with the following steps:*

(a) Plot the returns.

(b) Sorting the BOA weekly returns according to the order of the C weekly returns.

(c) Apply locally linear quantile regression on BOA and C: First regressing BOA on the ranks $\{i/n\}_{i=1}^{546}$*, and then divide it with the square root of the empirical pdf of C.*

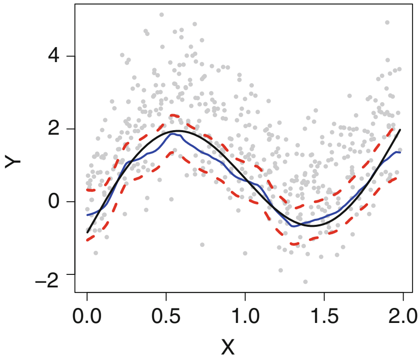


Fig. 16.19 Uniform Confidence Bands for $\tau = 0.1$ Expectile Curve. Theoretical Expectile Curve, Estimated Expectile Curve and 95 % Uniform Confidence Bands.
SFSconfexpectile0.95

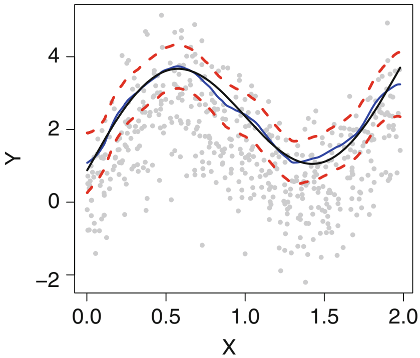


Fig. 16.20 Uniform Confidence Bands for $\tau = 0.9$ Expectile Curve. Theoretical Expectile Curve, Estimated Expectile Curve and 95 % Uniform Confidence Bands.
SFSconfexpectile0.95

(d) *Construct the confidence band for the quantile curve you get in the last problem with the theorem:*

Let $h = n^{-\delta}$, $\frac{1}{5} < \delta < \frac{1}{3}$, $\lambda(K) = \int_{-A}^{A} K^2(u)du$, *where* $K(\cdot)$ *is supported on* $[-A, A]$. $J = [0, 1]$. *Define* $c_1(K) = \{K^2(A) + K^2(-A)\}/2\lambda(K)$, $c_2(K) = \int_{-A}^{A}\{K'(u)\}^2 du/2\lambda(K)$ *and*

$$d_n = \begin{cases} (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2}\Big[\log\frac{\{c_1(K)}{\pi^{1/2}}\} \\ +\frac{1}{2}\{\log\delta + \log\log n\}\Big], \ \textit{if}\ c_1(K) > 0; \\ \\ (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2}\log\frac{\{c_2(K)}{2\pi}\}, \\ \textit{otherwise}. \end{cases}$$

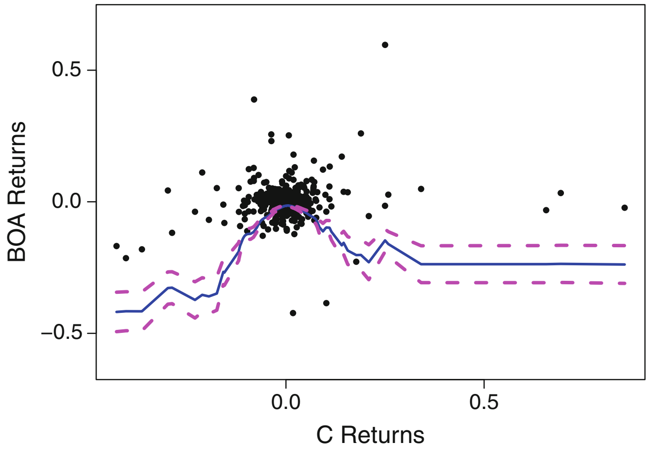


Fig. 16.21 The $\tau = 5\,\%$ quantile curve (*solid line*) and its 95 % confidence band (*dashed line*). SFSbootband

Then

$$\mathrm{P}\left[(2\delta\log n)^{1/2}\left\{\sup_{x\in J} r(x)\frac{|l_n(x)-l(x)|}{\lambda(K)^{1/2}} - d_n\right\} < z\right] \to \exp\{-2\exp(-z)\},$$

as $n \to \infty$*, with*

$$r(x) = (nh)^{1/2} f\{l(x)|x\}\{f_X(x)/\tau(1-\tau)\}^{1/2},$$

where $f_X(\cdot)$ *is the marginal pdf for* X *and* $f(\cdot|x)$ *is the conditional pdf of* Y *on* $X = x$.

With the steps described, bandwidth $h = 0.2155$ specified can easily be obtained employing the R command “lprq” in the package “quantreg”, so we can produce the quantile curve. Applying the given theorem, we construct the 95 % confidence band. The outcome can be seen in Fig. 16.21. In the figure, the slopes are different on both tails. While on the right the slope is almost zero, it is quite positive on the left.

Exercise 16.17 (Bootstrap). *The confidence band we discussed in the last exercise can also be constructed via bootstrap, with the following steps:*

(a) *We have two asset returns sequence* $\{Z_i\}_{i=1}^n$ *and* $\{Y_i\}_{i=1}^n$. $\{X_i\}_{i=1}^n$*:* n *equally divided grid on* $[0,1]$. $n = 546$. *Assume that* Z *is ordered by size and* Y *has been sorted by the order of* Z.
(b) *Bivariate data:* $\{(X_i, Y_i)\}_{i=1}^n$ *Compute* $l_h(x)$ *of* $Y_1, \ldots, Y_n$ *and residuals* $\hat{\varepsilon}_i = Y_i - l_h(X_i),\ i = 1, \ldots, n$. $\tau = 5\,\%$. *The bandwidth is 0.2155(BOA-C)*

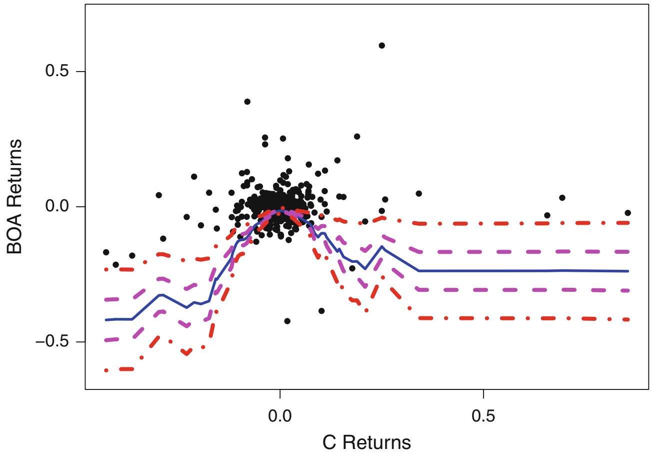


Fig. 16.22 The $\tau = 5\,\%$ quantile curve (*solid line*), 95 % confidence band (*dashed line*) and the bootstrapping 95 % confidence band (*dashed-dot line*). SFSbootband

(c) *Compute the conditional edf:*

$$\hat{F}(t|x) = \frac{\sum_{i=1}^{n} K_h(x - X_i)\mathbf{1}\{\hat{\varepsilon}_i \leqslant t\}}{\sum_{i=1}^{n} K_h(x - X_i)}$$

with the quartic kernel

$$K(u) = \frac{15}{16}(1 - u^2)^2, \ (|u| \leqslant 1).$$

(d) *Generate rv* $\varepsilon^*_{i,b} \sim \hat{F}(t|x)$, $b = 1, \ldots, B$ *and construct the bootstrap sample* $Y^*_{i,b}, i = 1, \ldots, n, b = 1, \ldots, B,\ B = 500$, *as follows:*

$$Y^*_{i,b} = l_g(X_i) + \varepsilon^*_{i,b},$$

with g=0.3774 (BOA-C).

(e) *For each bootstrap sample* $\{(X_i, Y^*_{i,b})\}_{i=1}^n$, *compute* l^*_h *and the random variable*

$$d_b \stackrel{\text{def}}{=} \sup_{x \in J^*} \left[\hat{f}\{l^*_h(x)|x\} \sqrt{\hat{f}_X(x)} |l^*_h(x) - l_g(x)| \right]. \tag{16.8}$$

(f) *Calculate the* $(1-\alpha)$ *quantile* d^*_α *of* $d_1, \ldots, d_B$.

(g) *Construct the bootstrap uniform confidence band centered around* $l(z) = l_h(x)/\sqrt{\hat{f}_Z(z)}$, *i.e.*

$$l(z) \pm \left[\hat{f}\{l_h(x)|x\} \sqrt{\hat{f}_X(x) \hat{f}_Z(z)} \right]^{-1} d^*_\alpha$$

.

The outcome of the bootstrapping 95 % confidence band can be seen in Fig. 16.22. The bootstrapping confidence band is wider than the asymptotic confidence band.

Chapter 17
Volatility Risk of Option Portfolios

Put all your eggs in one basket – and watch that basket.
Mark Twain, The Tragedy of Pudd'nhead Wilson.

There is a close connection between the value of an option and the volatility process of the financial underlying. Assuming that the price process follows a geometric Brownian motion, we have derived the Black-Scholes formula (BS) for pricing European options. With this formula and when the following values are given, the option price is, at a given time point, a function of the volatility parameters: τ (time to maturity in years), K (strike price), r (risk free, long-run interest rate) and S (the spot price of the underlying).

Although the volatility parameter is a constant in the BS setting, we can estimate an implied volatility (IV) function from observed option market prices by inverting the BS formula. In doing so we find volatility changes over time and moneyness, usually presented in a dynamic volatility smile surface.

By observing the volatility surface over time, distinct changes in the location and structure become obvious. Identifying the intertemporal dynamics is of central importance for a number of applications, such as the risk management of option portfolios.

Exercise 17.1 (Interpolation Methods). *On July 1st, 2005 the closing price of DAX was* 4,617.07. *One observes the following call options with strike* $K = 4{,}600$ *and maturities in year* $\tau_1 = 0.2109$, $\tau_2 = 0.4602$, $\tau_3 = 0.7095$. *The prices of these options are* $C_1 = 119.4$, $C_2 = 194.3$, $C_3 = 256.9$, *respectively. Assume that* C_2 *is not observed. In order to approximate this price one may use linear interpolation of options* C_1 *and* C_3. *The interpolation can be performed in prices* C *and in implied volatilities* $\widehat{\sigma}$.

Compare the two interpolation methods and check which gives the closest approximation to the true price. Interpolate also the variance and compare the results. (The interest rate is 2.1 %.*)*

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_17,

Using linear interpolation of C_1 and C_3, we have:

$$\begin{aligned}\frac{(C_3 - C_1)}{(\tau_3 - \tau_1)} &= \frac{(C_2 - C_1)}{(\tau_2 - \tau_1)} \\ C_2 &= \frac{(\tau_2 - \tau_1)(C_3 - C_1)}{(\tau_3 - \tau_1)} + C_1. \end{aligned} \tag{17.1}$$

Hence

$$\begin{aligned}\widehat{C_2^{(1)}} &= 188.15 \\ |\widehat{C_2^{(1)}} - C_2| &= |194.3 - 188.15| \\ &= 6.15.\end{aligned}$$

For the second approach we calculate the IV from options by inverting the BS formula:

$$C(S,\tau) = \exp\{(b-r)\tau\}S\Phi(y + \sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(y).$$

The IVs at τ_1 and τ_3 are estimated as $\widehat{\sigma}_1 = 0.1182$ and $\widehat{\sigma}_3 = 0.1377$.

As in (17.1) we interpolate the implied volatilities to get $\widehat{\sigma}_2$ and plug the result into the BS formula:

$$\begin{aligned}\widehat{\sigma_2} &= \frac{(\tau_2 - \tau_1)(\sigma_3 - \sigma_1)}{(\tau_3 - \tau_1)} + \sigma_1 \\ \widehat{\sigma_2} &= 0.128 \\ \widehat{C_2^{(2)}} &= 191.35 \\ |\widehat{C_2^{(2)}} - C_2| &= |194.3 - 191.35| \\ &= 2.95.\end{aligned}$$

We see that $6.15 > 2.95$ and conclude that the volatility interpolation approach leads to a more accurate approximation.

For the variance, we have as in (17.1):

$$\begin{aligned}\widehat{\sigma_2^2} &= \frac{(\tau_2 - \tau_1)(\sigma_3^2 - \sigma_1^2)}{(\tau_3 - \tau_1)} + \sigma_1^2 \\ \widehat{\sigma_2^2} &= 0.0165 \\ \widehat{C_2^{(3)}} &= 191.8\end{aligned}$$

$$\begin{aligned}|\widehat{C_2^{(3)}} - C_2| &= |194.3 - 191.8|\\ &= 2.5\end{aligned}$$

Using the interpolated variance gives a more accurate estimator in this case. SFSinterpolMaturity

Exercise 17.2 (Interpolation Methods and Approximation). *On July 1st, 2005 the closing price of DAX was* 4,617.07. *One observes the following call options with strikes* $K_1 = 4{,}000$, $K_2 = 4{,}200$, $K_3 = 4{,}500$ *and maturity in years* $\tau = 0.2109$. *The prices of these options are* $C_1 = 640.6$, $C_2 = 448.7$ $C_3 = 188.5$ *respectively. Assume that* C_2 *is not observed. In order to approximate this price one may use linear interpolation of options* C_1 *and* C_3. *The interpolation can be performed in prices* C *or in implied volatilities* $\widehat{\sigma}$.

Compare these interpolation methods and check which gives the closest approximation to the true price. Use interest rate $r = 2.1\,\%$.

Using the linear interpolation (17.1) of the prices C_1 and C_3 gives an approximation of C_2 equal to 459.76. The difference to the true price is 11.06.

Calculating implied volatilities from options with strike prices K_1 and K_3 yields $\widehat{\sigma}_1 = 0.1840$ and $\widehat{\sigma}_3 = 0.1276$. Interpolating the volatilities and plugging the result back into the Black-Scholes formula give an approximation of C_2 equal to 449.33. The difference to the true price is 0.63.

The interpolation of variances yields $\widehat{\sigma_2^2} = 0.0268$ and the approximation for C_2 equals to 450.11. The difference to the true price is 1.41. Hence the interpolated volatility gives the best approximation in this case.

SFSinterpolStrike

Exercise 17.3 (Stickiness). *Let the current underlying price be* $S_0 = 100$, *maturity* $\tau = 0.25$ *years and interest rate* $r = 2\,\%$. *Assume that implied volatility is given as function of strike price* $f(K) = 0.000167K^2 - 0.03645K + 2.08$.

Plot call option prices as a function of strikes for $K \in (85, 115)$. *Assume that the underlying price moves to* $S_1 = 105$. *The implied volatility function may be fixed to the strike prices (sticky strike) or moneyness* K/S_1 *(sticky moneyness). Plot call option prices with two different stickiness assumptions. Compare the relative difference of both approaches.*

For the calculation of the call prices the Black-Scholes formula needs to be applied with the given inputs. In case the underlying price shifts to $S_1 = 105$ and the sticky strike assumption, only the spot price has to be updated. For the sticky moneyness assumption the function of implied volatility must be scaled by S_0/S_1 i.e.

$$\widetilde{f}(K) = f(KS_0/S_1).$$

The call prices are shown in Fig. 17.1. The relative difference between two stickiness assumption is displayed in Fig. 17.2. Note that it is negligible for the

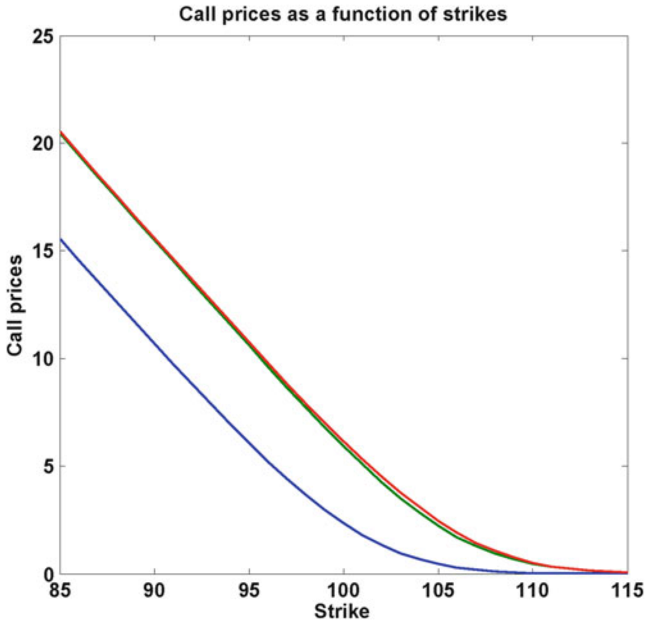


Fig. 17.1 Call prices as a function of strikes for $r = 2\,\%$, $\tau = 0.25$. The implied volatility functions curves are given as $f(K) = 0.000167K^2 - 0.03645K + 2.08$ (*blue* and *green curves*) and $\widetilde{f}(K) = f(KS_0/S_1)$ (*red curve*). The level of underlying price is $S_0 = 100$ (*blue*) and $S_1 = 105$ (*green*, *red*) SFSstickycall

in-the-money options and is significant for the out-of-the-money options. Additionally, Fig. 17.3 presents the implied volatility smile as a function of strike prices and the smile is obtained for the sticky moneyness assumption. The function shifts to the right when the underlying jumps up.

Exercise 17.4 (Risk Reversal). *A risk reversal strategy is defined as a long position in an out-of-the-money put and a short position in an out-of-the-money call (or vice versa). Consider the risk reversal strategy of long put with strike* 85 *and short call with strike* 115 *for maturity* $\tau = 0.25$ *years. Let the current underlying price be* $S_0 = 100$*, and interest rate* $r = 2\,\%$*. Compare the prices of the risk reversal for the following implied volatility curves given as a function of strike price:*

$$f_1(K) = 0.000167K^2 - 0.03645K + 2.080,$$
$$f_2(K) = 0.000167K^2 - 0.03645K + 2.090,$$
$$f_3(K) = 0.000167K^2 - 0.03517K + 1.952.$$

In order to calculate the price of the risk reversal, the spot price, strike, interest rate, maturity and volatility need to be plugged into the BS formula. For the calculation of the volatilities, the given functions have to be evaluated at $K = 85$ and $K = 115$. Hence, by using the BS formula the prices of the risk reversal strategy for the functions f_1, f_2 and f_3 are 0.1260, 0.1603 and 0.0481, respectively.

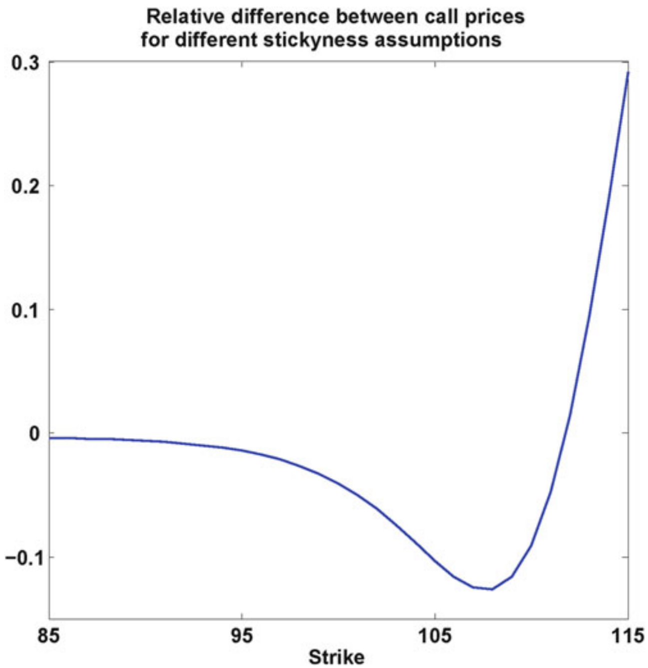


Fig. 17.2 Relative differences of the call prices for two different stickiness assumptions SFSstickycall

The considered functions are displayed in Fig. 17.4. The panels compare the implied volatility functions f_2 and f_3 to the function f_1. It can be recognized that f_2 represents a parallel shift of f_1, while f_3 tilts f_1. Vega, defined as $\frac{\partial C}{\partial \sigma}$, i.e. price change with respect to volatility, for plain vanilla options is positive. Therefore both the value of the long put and of the short call increase. The aggregate change is not zero due to the difference in the vega of the two options. The tilting of the volatility curve implies different direction of volatility changes for the considered options. As $K < S_0$, $f_3 < f_1$ and as $K > S_0$, $f_3 > f_1$, hence the values of long put and short call both decrease. Therefore, one would expect skew changes to have a bigger impact on the risk reversals than a parallel shift.

Exercise 17.5 (Calendar Spread). *A calendar spread strategy is defined as a position in two options with same strike but different maturity. Consider a calendar spread for an at-the-money short call with maturity $\tau_1 = 0.25$ and an at-the-money long call with maturity $\tau_2 = 1$ year. Let the current underlying price be $S_0 = 100$, and interest rate $r = 2\,\%$. Compare the prices of the calendar spread for the following implied volatility curves given as functions of maturity:*

$$f_1(\tau) = 0.15\tau + 0.05,$$
$$f_2(\tau) = 0.15\tau + 0.06,$$
$$f_3(\tau) = 0.1\tau + 0.075.$$

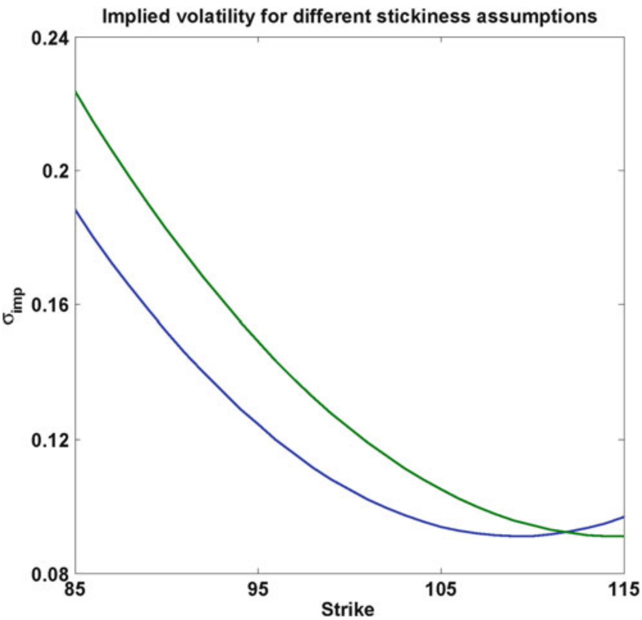


Fig. 17.3 Implied volatility functions $f(K) = 0.000167K^2 - 0.03645K + 2.08$ and $\widetilde{f}(K) = f(KS_0/S_1)$ SFSstickycall

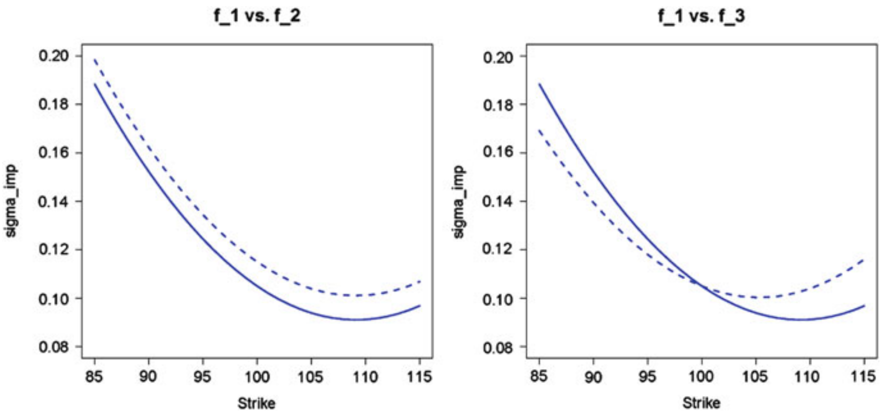


Fig. 17.4 The implied volatility functions f_1, f_2 and f_3. *Left panel*: comparison of f_1 (*solid line*) and f_2 (*dashed line*). *Right panel*: comparison of f_1 (*solid line*) and f_3 (*dashed line*) SFSriskreversal

In order to calculate the price of the calendar spread the spot price, strike, interest rate, maturity and volatility need to be plugged into the BS formula. For the calculation of the volatilities the given functions have to be evaluated in $\tau = 0.25$ and $\tau = 1$. Hence the prices of risk reversal for functions f_1, f_2 and f_3 are 6.9144, 7.1077 and 5.6897, respectively.

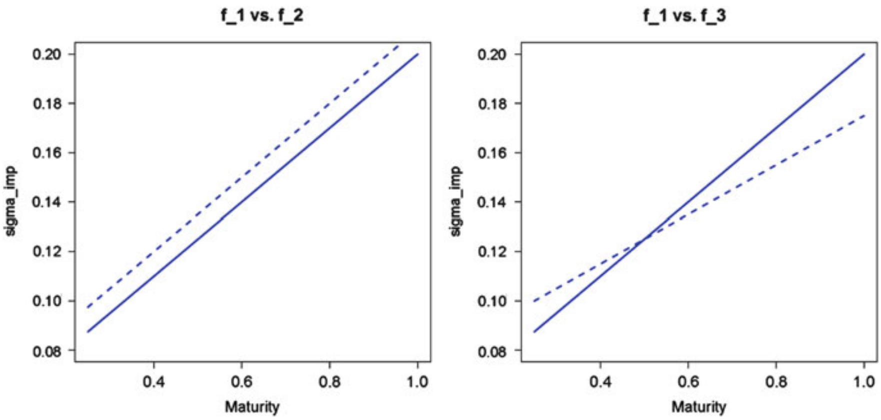


Fig. 17.5 The implied volatility functions f_1, f_2 and f_3. *Left panel*: comparison of f_1 and f_2. *Right panel*: comparison of f_1 and f_3 `SFScalendarspread`

Table 17.1 Observed strikes, implied volatilities and call prices

Observation	K	σ	C_t
1	4,000	0.1840	640.6
2	4,100	0.1714	543.8
3	4,200	0.1595	448.7
4	4,500	0.1275	188.5

The considered functions are displayed in Fig. 17.5. The panels compare the functions f_2 and f_3 to the function f_1. It can be recognized that f_2 represents the parallel shift of f_1, while f_3 tilts f_1. The upward shift of the volatility curve triggers an opposite change in the price of the calendar spread. The value of the short call decreases the value, while the change in value of the long call increases the price of the calendar spread. Due to the increasing vega and call price in time maturity the overall upward shift results in increase of the calendar spread value. The tilting of the volatility term structure implies the common movement in the calendar spread's components. The values of the short call and of the long call decrease when the curve flattens.

Exercise 17.6 (Implied Volatility). *In order to price options for strikes outside the observed range an extrapolation has to be used. Consider the IV data given in Table 17.1*

Let the observation 1 *be a validation observation. Apply constant extrapolation of the IV on observation* 2*, linear extrapolation of the IV on observations* 2 *and* 3*, and quadratic extrapolation of the IV on observations* $2, 3$ *and* 4 *to obtain an estimate of the call price for observation* 1*. Compare the results with the actual price. How would the results differ if instead of extrapolating in IVs the true price would be used? For your calculation use spot* $S_0 = 4{,}617.07$*, interest rate* $r = 2.1\,\%$ *and maturity* $\tau = 0.2109$.

Constant extrapolation of IV would assume the given volatility value i.e. 0.1714, the linear extrapolation would lead to 0.1833 and quadratic extrapolation to 0.1839. The call prices obtained from these volatility estimates are 638.68, 640.48 and 640.58, respectively. Using the extrapolation methods based on the observed call prices leads to the following approximations: 543.80, 638.90, 643.08. Note that extrapolation performed in volatilities leads to smaller errors than the extrapolation performed directly on the prices.

`SFSextrapolationIV`

Chapter 18
Portfolio Credit Risk

Winning is earning, losing is learning.

Financial institutions are interested in loss protection and loan insurance. Thus determining the loss reserves needed to cover the risk stemming from credit portfolios is a major issue in banking. By charging risk premiums a bank can create a loss reserve account which it can exploit to be shielded against losses from defaulted debt. However, it is imperative that these premiums are appropriate to the issued loans and to the credit portfolio risk inherent to the bank. To determine the current risk exposure it is necessary that financial institutions can model the default probabilities for their portfolios of credit instruments appropriately. To begin with, these probabilities can be viewed as independent but it is apparent that it is plausible to drop this assumption and to model possible defaults as correlated events.

In this chapter we give examples of the different methods to calculate the risk exposure of possible defaults in credit portfolios. Starting with basic exercises to determine the loss given default and the default probabilities in portfolios with independent defaults, we move on to possibilities to model correlated defaults by means of the Bernoulli and Poisson mixture models.

Exercise 18.1 (Expected Loss). *Assume a zero coupon bond repaying full par value 100 with probability 95 % and paying 40 with probability 5 % in 1 year. Calculate the expected loss.*

Probability of default in this exercise is $PD = 5\,\%$, exposure at default (EAD) is $EAD = 100$ and loss given default (LGD) is $LGD = 60\,\%$. Hence, the expected loss is:

$$\mathsf{E}(\widetilde{L}) = EAD \cdot LGD \cdot PD = 100 \times 0.6 \times 0.05 = 3$$

Exercise 18.2 (Expected Loss). *Consider a bond with the following amortization schedule: the bond pays 50 after half a year (T_1) and 50 after a full year (T_2). In case of default before T_1 the bond pays 40 and in case of default in $[T_1, T_2]$ pays 20.*

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5_18,

Calculate the expected loss when the probabilities of default in $[0, T_1)$ *and* $[T_1, T_2]$ *are*

(a) 1 and 4 %
(b) 2.5 and 2.5 %
(c) 4 and 1 %

respectively.

Following the expected loss logic (Exercise 18.1) one obtains

$$(a)\ \mathsf{E}(\widetilde{L}) = 60 \times 0.010 + 30 \times 0.040 = 0.6 + 1.20 = 1.80$$
$$(b)\ \mathsf{E}(\widetilde{L}) = 60 \times 0.025 + 30 \times 0.025 = 1.5 + 0.75 = 2.25$$
$$(c)\ \mathsf{E}(\widetilde{L}) = 60 \times 0.040 + 30 \times 0.010 = 2.4 + 0.30 = 2.70.$$

Note that the time of default has an impact on the expected loss. Front loaded default curves generate a larger expected loss than back loaded curves.

Exercise 18.3 (Joint Default). *Consider a simplified portfolio of two zero coupon bonds with the same probability of default (PD), par value 1 and 0 recovery. The loss events are correlated with correlation* ρ*.*

(a) Calculate the loss distribution of the portfolio,
(b) Plot the loss distribution for PD = 20 % and $\rho = 0; 0.2; 0.5; 1$.

(a) Let L_1 and L_2 be the loss of the first and second bond respectively. Then

$$\begin{aligned}\mathsf{Corr}(L_1, L_2) &= \frac{\mathsf{Cov}(L_1, L_2)}{\sqrt{\mathsf{Var}(L_1)\,\mathsf{Var}(L_2)}} \\ &= \frac{\mathsf{E}(L_1 L_2) - \mathsf{E}(L_1)\,\mathsf{E}(L_2)}{\mathsf{Var}\,L_1} \\ &= \frac{\mathrm{P}(L_1 = 1, L_2 = 1) - PD^2}{(1 - PD)PD}\end{aligned}$$

and

$$\mathrm{P}(L_1 = 1, L_2 = 1) = \rho(1 - PD)PD + PD^2.$$

Note that for $\rho = 0$, i.e. the losses are uncorrelated, the joint probability is equal to PD^2. For $\rho = 1$ they are linearly dependent and the joint probability is equal to PD.

$$\mathrm{P}(L_1 = 1, L_2 = 0) + \mathrm{P}(L_1 = 1, L_2 = 1) = \mathrm{P}(L_1 = 1) = PD$$

and hence

$$\mathrm{P}(L_1 = 1, L_2 = 0) = PD - \rho(1 - PD)PD - PD^2 = PD(1 - PD)(1 - \rho).$$

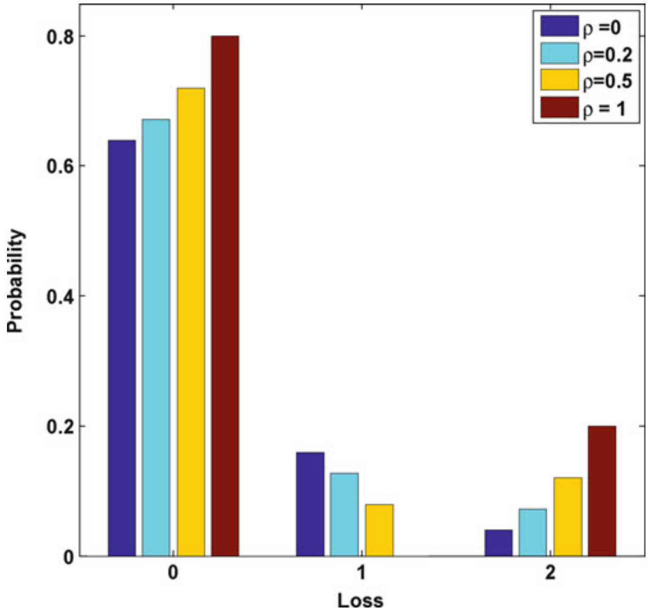


Fig. 18.1 The loss distribution of the two identical losses with probability of default 20 % and different levels of correlation i.e. $\rho = 0, 0.2, 0.5, 1$ SFSLossDiscrete

In case of independent losses, the probability that only one bond defaults is equal to $PD(1 - PD)$ and for fully dependent bonds it reduces to zero as they jointly behave as one asset.

$$P(L_1 = 0, L_2 = 0) + P(L_1 = 1, L_2 = 0) = P(L_2 = 0) = 1 - PD$$

and

$$P(L_1 = 0, L_2 = 0) = (1 - PD) - PD(1 - PD)(1 - \rho)$$

For $\rho = 0$ the formula reduces to $(1 - PD)^2$ and for $\rho = 1$ it is as expected equal to $(1 - PD)$.

From these calculations the resulting loss distribution of $L = L_1 + L_2$ is given by:

$$P(L = 2) = \rho(1 - PD)PD + PD^2$$
$$P(L = 1) = 2PD(1 - PD)(1 - \rho)$$
$$P(L = 0) = (1 - PD) - PD(1 - PD)(1 - \rho).$$

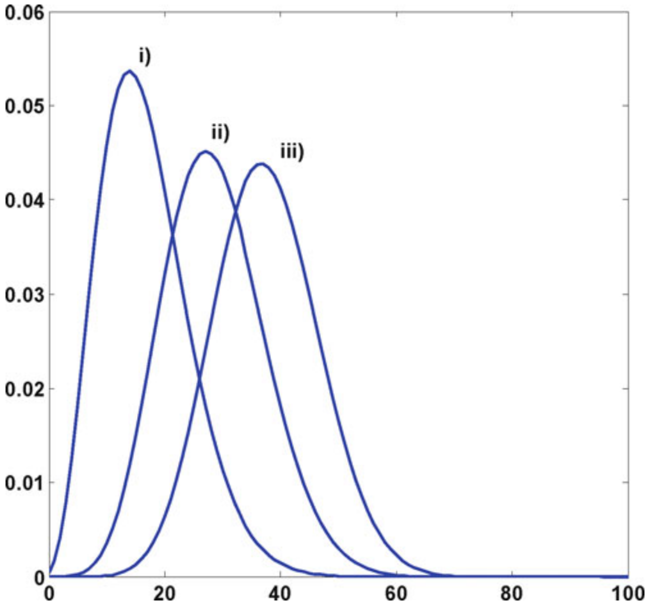


Fig. 18.2 Loss distribution in the simplified Bernoulli model. Presentation for cases (i)–(iii). Note that for visual convenience a solid line is displayed although the true distribution is a discrete distribution `SFSLossBern`

(b) See Fig. 18.1. While the correlation increases from 0 to 1, the probability of having only one loss tends to zero and the probabilities of no loss and two losses increases. This logic is also presented for the continuous case.

Exercise 18.4 (Bernoulli Model). *Consider a simplified Bernoulli model of $m = 100$ homogeneous risks with the same loss probabilities P_i, P coming from the beta distribution. The density of the beta distribution is*

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\mathbf{1}\{x \in (0,1)\}.$$

Plot the loss distribution of $L = \sum_{i=1}^{m} L_i$ for the following set of parameters

$$\begin{aligned}
&(i)\ \ \alpha = 5,\ \ \beta = 25\\
&(ii)\ \ \alpha = 10, \beta = 25\\
&(iii)\ \alpha = 15, \beta = 25\\
&(iv)\ \alpha = 5,\ \ \beta = 45
\end{aligned}$$

$$(v) \quad \alpha = 10, \beta = 90$$
$$(vi) \quad \alpha = 20, \beta = 180$$

Given P, the L_i are independent and $\mathrm{P}(L = k|\mathrm{P} = p) = \binom{m}{k} p^k (1-p)^{(m-k)}$. To obtain the unconditional distribution one simply needs to integrate with respect to the mixing distribution

$$\mathrm{P}(L = k) = \int_0^1 \binom{m}{k} p^k (1-p)^{(m-k)} f(p) dp$$

Note that changing α allows for adjusting the expected loss, cases (i)–(iii) (See Fig. 18.2) Figure 18.3 presents the situation when the expected loss stays constant and the distributions have different variances, cases (iv)–(vi).

Exercise 18.5 (Poisson Model). *Consider a simplified Poisson model of* $m = 100$ *homogeneous risks with same intensities* $\Lambda_i = \Lambda$ *coming from the gamma distribution. The density of the gamma distribution is*

$$f(x) = \{\Gamma(\alpha)\beta^\alpha\}^{-1} x^{\alpha-1} \exp(-x/\beta).$$

Plot the loss distribution of $L = \sum_{i=1}^m L_i$ *for the following set of parameters*

$$(i) \quad \alpha = 2, \quad \beta = 5$$
$$(ii) \quad \alpha = 4, \quad \beta = 5$$
$$(iii) \quad \alpha = 6, \quad \beta = 5$$
$$(iv) \quad \alpha = 3, \quad \beta = 3.33$$
$$(v) \quad \alpha = 2, \quad \beta = 5$$
$$(vi) \quad \alpha = 10, \beta = 1$$

Given Λ, L_i are independent and

$$\mathrm{P}(L = k|\Lambda = \lambda) = \frac{\exp(-m\lambda)(m\lambda)^k}{k!}.$$

The unconditional distribution is obtained by

$$P(L = k|\Lambda = \lambda) = \int_0^{+\infty} \frac{\exp(-m\lambda)(m\lambda)^k}{k!} f(\lambda) d\lambda.$$

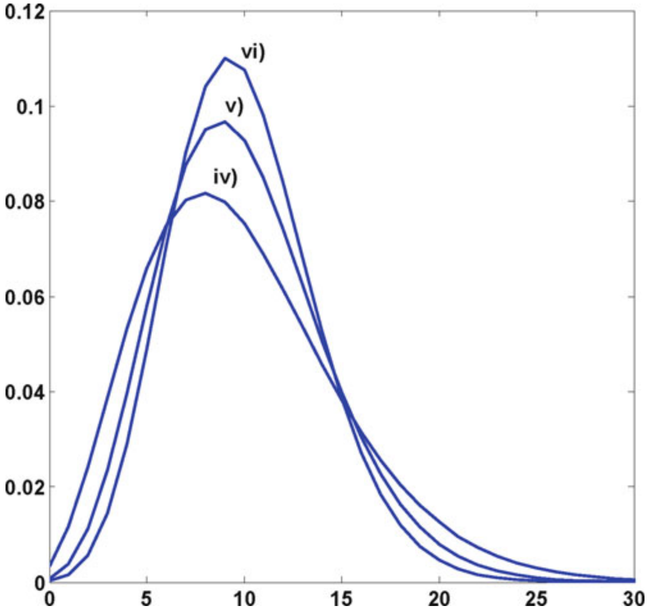


Fig. 18.3 Loss distribution in the simplified Bernoulli model. Presentation for cases (iv)–(vi). Note that for the visual convenience a solid line is displayed although the true distribution is a discrete distribution `SFSLossBern`

It is easy to observe that α allows for adjusting the expected loss, cases (i)–(iii), as displayed in Fig. 18.4. Figure 18.5 presents the situation when the expected loss stays constant and the distributions have different variances, cases (iv)–(vi).

Exercise 18.6 (Bernoulli vs. Poisson Model). *Consider the Bernoulli model with the same loss probabilities $P_i = P$ and the Poisson model with intensities $\Lambda_i = \Lambda$. Assume that P and Λ have the same mean and variance.*

(a) Show that the variance of the individual loss in the Poisson model exceeds the variance of the individual loss in the Bernoulli model.
(b) Show that the correlation of two losses in the Poisson model is smaller than in the Bernoulli model.

(a) In the Bernoulli model

$$\begin{aligned}\mathsf{Var}(L_i) &= \mathsf{E}(P_i)\{1 - \mathsf{E}(P_i)\}\\ &= \mathsf{E}(\Lambda_i)\{1 - \mathsf{E}(\Lambda_i)\}.\end{aligned}$$

In the Poisson model $\mathsf{Var}(L_i) = \mathsf{E}(\Lambda_i) + \mathsf{Var}(\Lambda_i)$, which is clearly greater.

(b) This fact is implied by (a) since in the Poisson model the denominator in the correlation formula is greater.

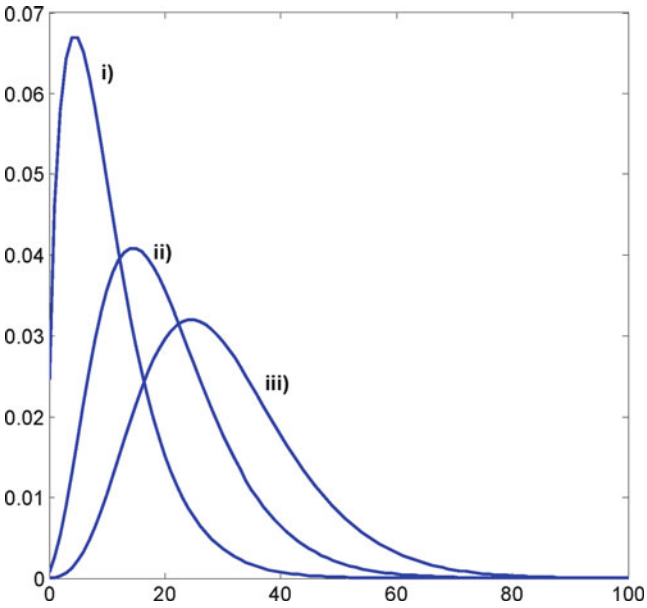


Fig. 18.4 Loss distribution in the simplified Poisson model. Presentation for cases (i)–(iii). Note that for visual convenience a solid line is displayed although the true distribution is a discrete distribution SFSLossPois

Exercise 18.7 (Moments, Correlation and Tail Behaviour of Bernoulli and Poisson Model). *Consider the Bernoulli model from Exercise 18.4 with* $\alpha^B = 1$, $\beta^B = 9$ *and the Poisson model from Exercise 18.5 with* $\alpha^P = 1.25$, $\beta^P = 0.08$.

(a) Show that the cumulative loss distributions have same first two moments.
(b) Calculate $\mathsf{Corr}(L_i, L_j)$ *for these two models.*
(c) Plot both densities in one figure and discuss their tail behavior.

(a) Bernoulli distribution

$$\mathsf{E}(L) = \sum_{i=1}^{m} \mathsf{E}(L_i) = m\,\mathsf{E}(P) = m\frac{\alpha^B}{\alpha^B + \beta^B} = 0.1 \cdot m$$

$$\begin{aligned}
\mathsf{Var}(L) &= \mathsf{Var}\{\mathsf{E}(L|P)\} + \mathsf{E}\{\mathsf{Var}(L|P)\} \\
&= \mathsf{Var}(mP) + \mathsf{E}\{mP(1-P)\} \\
&= m^2\,\mathsf{Var}(P) + m\,\mathsf{E}(P) - m\,\mathsf{E}(P^2) \\
&= (m^2 - m)\,\mathsf{Var}(P) + m\,\mathsf{E}(P) - m\,\mathsf{E}(P)^2 \\
&= (100^2 - 100)\cdot\frac{9}{10\cdot 10\cdot 11} + 100\cdot 0.1 - 100\cdot 0.1^2 \\
&= 81 + 10 - 1 = 90
\end{aligned}$$

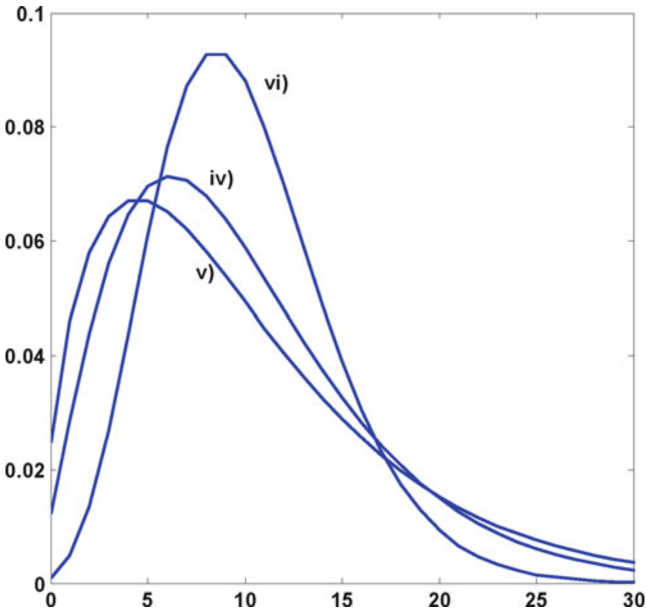


Fig. 18.5 Loss distribution in the simplified Poisson model. Presentation for cases (iv)–(vi). Note that for the visual convenience the solid line is displayed although the true distribution is a discrete distribution SFSLossPois

Poisson distribution

$$\mathsf{E}(L) = \sum_{i=1}^{m} \mathsf{E}(L_i) = m\,\mathsf{E}(\Lambda) = m\alpha^P\beta^P = 0.1 \cdot m$$

$$\begin{aligned}\mathsf{Var}(L) &= \mathsf{Var}\{\mathsf{E}(L|\Lambda)\} + \mathsf{E}\{\mathsf{Var}(L|\Lambda)\}\\ &= \mathsf{Var}(m\Lambda) + \mathsf{E}\{m\Lambda\}\\ &= m^2\,\mathsf{Var}(\Lambda) + m\,\mathsf{E}(\Lambda)\\ &= 100^2 \cdot 1.25 \cdot 0.08^2 + 100 \cdot 0.1\\ &= 80 + 10 = 90\end{aligned}$$

(b) Bernoulli distribution

$$\mathsf{Corr}(L_i, L_j) = \frac{\mathsf{Var}(P)}{\mathsf{E}(P)\{1 - \mathsf{E}(P)\}} = \frac{0.0082}{0.1 \cdot 0.9} = 0.0909$$

Poisson distribution

$$\mathsf{Corr}(L_i, L_j) = \frac{\mathsf{Var}(\Lambda)}{\mathsf{Var}(\Lambda) + \mathsf{E}(\Lambda)} = \frac{1.25 \cdot 0.08^2}{1.25 \cdot 0.08^2 + 0.1} = 0.0741$$

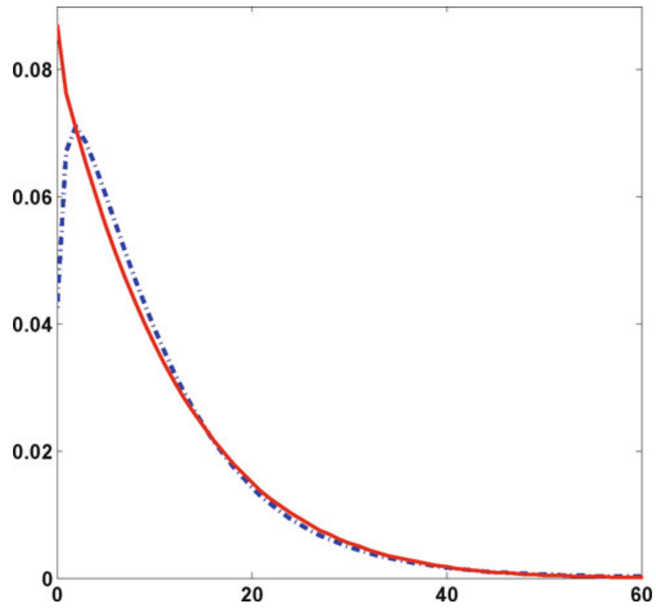


Fig. 18.6 Loss distributions in the simplified Bernoulli model (*straight line*) and simplified Poisson model (*dotted line*) SFSLossBernPois

(c) From Exercise 18.6 we know that there is a systematic difference between the Bernoulli and Poisson model. Even if the first and second moments of the two distributions match, the variance in the Poisson model will always be greater than the variance of the Bernoulli model. This effect evidently leads to lower default correlations in the Poisson model. Lower default correlations in the loss distribution will result in thinner tails and vice versa. This is shown in Figs. 18.6 and 18.7.

Exercise 18.8 (Calibration of representative Portfolio). *Assume a portfolio of N obligors. Each asset has a notional value EAD_i, probability of default p_i, correlation between default indicators $\mathsf{Corr}(L_i, L_j) = \rho_{i,j}$ for $i, j = 1, \dots, N$. For simplicity assume no recovery. Analysis of the loss distribution of this portfolio can be simplified by assuming a homogeneous portfolio of D uncorrelated risks with same notional $\widetilde{EAD}$ and probability of default $\widetilde{p}$. Calibrate the representative portfolio such that total exposure, expected loss and variance match the original portfolio.*

In order to match the exposure one obtains the following:

$$\sum_{i=1}^{N} EAD_i = D\widetilde{EAD}.$$

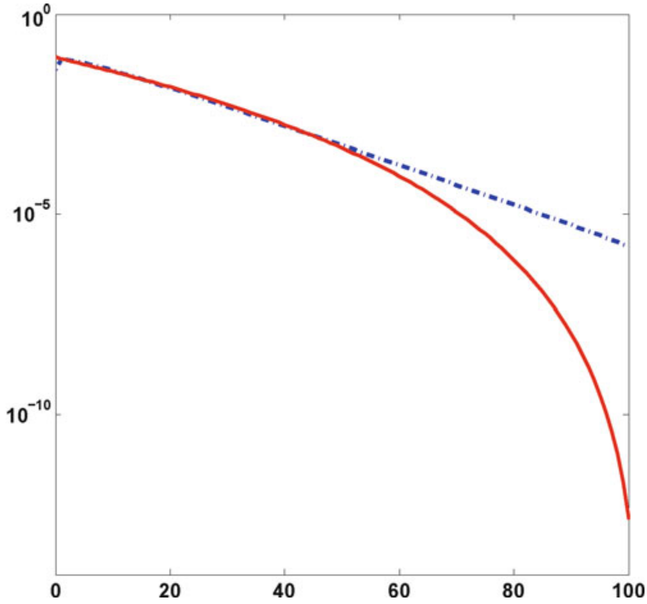


Fig. 18.7 The higher default correlations result in fatter tails of the simplified Bernoulli model (*straight line*) in comparison to the simplified Poisson model (*dotted line*) SFSLossBernPois

Matching expected loss gives:

$$\mathsf{E}\,\widetilde{L} = \sum_{i=1}^{N} EAD_i p_i = D\widetilde{EAD}\widetilde{p},$$

which leads to

$$\widetilde{p} = \frac{\sum_{i=1}^{N} EAD_i p_i}{\sum_{i=1}^{N} EAD_i}.$$

$\widetilde{p}$ is weighted by the notional probability of all obligors. The same variance requirement gives:

$$\begin{aligned}\mathsf{Var}(\widetilde{L}) = D\widetilde{EAD}^2\widetilde{p}(1-\widetilde{p}) &= \sum_{i=1}^{N}\sum_{j=1}^{N} EAD_i EAD_j \,\mathsf{Cov}(L_i, L_j)\\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} EAD_i EAD_j \sqrt{p_i(1-p_i)p_j(1-p_j)}\rho_{i,j}.\end{aligned}$$

Hence

$$\widetilde{EAD} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} EAD_i EAD_j \sqrt{p_i(1-p_i)p_j(1-p_j)}\rho_{i,j}}{\sum_{i=1}^{N} EAD_i \widetilde{p}(1-\widetilde{p})}$$

$$= \frac{\sum_{i=1}^{N} EAD_i \sum_{i=1}^{N}\sum_{j=1}^{N} EAD_i EAD_j \sqrt{p_i(1-p_i)p_j(1-p_j)}\rho_{i,j}}{\sum_{i=1}^{N} EAD_i p_i (\sum_{i=1}^{N} EAD_i - \sum_{i=1}^{N} EAD_i p_i)},$$

and

$$D = \frac{\sum_{i=1}^{N} EAD_i p_i (\sum_{i=1}^{N} EAD_i - \sum_{i=1}^{N} EAD_i p_i)}{\sum_{i=1}^{N}\sum_{j=1}^{N} EAD_i EAD_j \sqrt{p_i(1-p_i)p_j(1-p_j)}\rho_{i,j}}. \quad (18.1)$$

Exercise 18.9 (Homogeneous Portfolio). *Follow the assumption from Exercise 18.8 for the homogeneous portfolio i.e.* $EAD_i = EAD$, $p_i = p$, $\rho i, i = 1$ *for each* $i = 1, \ldots, N$ *and* $\rho_{i,j} = \rho$ *for* $i \neq j$. *Calculate the value of D for N=100, and* $\rho = 0, 2, 5, 10\,\%$.

Plugging EAD, p and ρ into the formula (18.1) one obtains:

$$D = \frac{N}{(N-1)\rho + 1}.$$

For the given correlation levels $\rho = 0, 2, 5, 10\,\%$ the values of D are $100, 33.5, 16.8, 9.2$. Since the loss distribution is approximated by the binomial distribution the values of D are rounded to the nearest integer number. Note that the increase in the correlation results in a decrease of the number of assets in the approximated uncorrelated portfolio.

References

Breiman, L. (1973). *Statistics: With a view towards application*. Boston: Houghton Mifflin Company.
Cizek, P., Härdle, W., & Weron, R. (2011). *Statistical tools in finance and insurance* (2nd ed.). Berlin/Heidelberg: Springer.
Feller, W. (1966). *An introduction to probability theory and its application* (Vol. 2). New York: Wiley.
Franke, J., Härdle, W., & Hafner, C. (2011). *Statistics of financial markets* (3rd ed.). Berlin/Heidelberg: Springer.
Härdle, W., & Simar, L. (2012). *Applied multivariate statistical analysis* (3rd ed.). Berlin: Springer.
Härdle, W., Müller, M., Sperlich, S., & Werwatz, A. (2004). *Nonparametric and semiparametric models*. Berlin: Springer.
Harville, D. A. (2001). *Matrix algebra: Exercises and solutions*. New York: Springer.
Klein, L. R. (1974). *A textbook of econometrics* (2nd ed., 488 p.). Englewood Cliffs: Prentice Hall.
MacKinnon, J. G. (1991). Critical values for cointegration tests. In R. F. Engle & C. W. J. Granger (Eds.), *Long-run economic relationships readings in cointegration* (pp. 266–277). New York: Oxford University Press.
Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. Duluth/London: Academic.
RiskMetrics. (1996). J.P. Morgan/Reuters (4th ed.). RiskMetricsTM.
Serfling, R. J. (2002). *Approximation theorems of mathematical statistics*. New York: Wiley.
Tsay, R. S. (2002). *Analysis of financial time series*. New York: Wiley.

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5, © Springer-Verlag Berlin Heidelberg 2013

Index

S. Borak et al., *Statistics of Financial Markets*, Universitext,
DOI 10.1007/978-3-642-33929-5, © Springer-Verlag Berlin Heidelberg 2013