PRINCETON SERIES IN FINANCE

THE QUANTITATIVE RISK MANAGEMENT EXERCISE BOOK

Marius Hofert, Rüdiger Frey and Alexander J. McNeil

The Quantitative Risk Management Exercise Book

Marius Hofert

Rüdiger Frey

Alexander J. McNeil

The Quantitative Risk Management Exercise Book

Marius Hofert

Rüdiger Frey

Alexander J. McNeil

Princeton University Press

Princeton and Oxford

Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
6 Oxford Street, Woodstock, Oxfordshire OX20 1TR

press.princeton.edu

ISBN 9780691206707

The publisher would like to acknowledge the authors for providing this manual in PDF form

To Saisai

To Catharina, Sebastian and Michaela

To Janine, Alexander and Calliope

Preface

Over the past fifteen years we have taught many courses on Quantitative Risk Management (QRM) to students, academics, financial practitioners and actuaries based on the textbook Quantitative Risk Management: Concepts, Techniques and Tools (McNeil, Frey and Embrechts; 2005, 2015). Henceforth we refer to this book as the *QRM textbook* and use the abbreviated citations MFE (2005) and MFE (2015) for the two editions.

We have used the QRM textbook to deliver lecture courses to students at our own academic institutions including ETH Zurich, Heriot-Watt University, the University of Leipzig, the Vienna University of Economics and Business (WU), the University of Washington (UW), the National University of Singapore (NUS) and the University of Waterloo. We are also aware that many colleagues have used the QRM textbook to teach courses at universities around the world.

The QRM textbook has also formed the basis for numerous intensive training courses, workshops and summer schools which have been delivered to financial practitioners and actuaries. In particular, it has been used as part of the CERA (Chartered Enterprise Risk Actuary) education programme of the European Actuarial Academy (EAA).

Throughout this time, we have been conscious that the QRM textbook has lacked a formal exercise collection to assist instructors and to provide students with the opportunity to test, consolidate and extend their knowledge. *The QRM Exercise Book* aims to fill this gap. It contains a comprehensive selection of the exercises that we have amassed during our own teaching experience, as well as a few exercises donated by colleagues.

The structure of the exercise book follows Chapters 1–11 of MFE (2015). Within chapters, exercises are typically grouped into three categories: review, basic and advanced questions. Review questions are designed to help readers to test their understanding of the concepts presented in the corresponding chapter of MFE (2015). As such, the solutions are usually found in MFE (2015) or involve at most simple calculations based on material in MFE (2015).

Basic questions are designed to test and enhance the understanding of the technical material. Typical examples might require students to calculate specific examples based on general techniques or might ask for simple proofs and minor extensions of results in MFE (2015). Taken together, review and basic questions are intended to be at the level of examination questions and should not require particularly long answers. Advanced exercises are either more lengthy, more difficult or take the reader somewhat beyond the textbook material.

Occasionally, a programming exercise is included. We recommend that these are tackled using the open source R language and environment for statistical computing. To facilitate the solution of programming exercises we provide references to publicly available R datasets and packages in the questions.

In view of the more qualitative and discursive nature of Chapter 1 of the QRM textbook, the exercises and solutions are structured in a slightly different way. Case study and discussion questions take the place of basic and advanced questions respectively. Case study questions invite students to research a number of well-known company failures, market crashes and natural catastrophes and consider their causes and/or their financial and economic implications. Discussion questions are more open-ended questions that address, for example, current trends and controversies in risk management.

The primary resource for solving these exercises is MFE (2015) although the exercise book also contains references to other books and papers that take the reader further in certain areas. An updated book of solutions to the exercises and a set of R scripts solving the programming exercises are available on the QRM Tutorial website `qrmtutorial.org`. For instructors there are some additional resources on the website including slides and extensive demonstration scripts in R.

We are grateful to Paul Embrechts who initiated the QRM project during his time as Professor of Insurance Mathematics at ETH Zurich. We have all benefited from his inspiration and mentoring over the years. We would also like to thank the Forschungsinstitut für Mathematik (FIM) at ETH Zurich for its financial support during multiple visits by the authors. Thanks are also due to several doctoral and postdoctoral teaching assistants in Zurich who helped develop the stock of examples and exercises for teaching QRM, including Valeria Bignozzi, Bikramjit Das, Catherine Donnelly, Edgars Jakobsons, Georg Mainik and Johanna G. Nešlehová, as well as to colleagues, friends and students around the world who have contributed exercises and solutions, in particular, Andrew Cairns and Jialing Han.

Contents

1 Risk in Perspective

Review

Exercise 1.1 (Types of financial risk)
Give definitions and examples for each of the following types of risk in the financial context:

a) market risk,
b) credit risk,
c) operational risk,
d) model risk, and
e) systemic risk.

Exercise 1.2 (Financial crisis of 2007–9)
Give an outline of the financial crisis of 2007–9 describing the main events and reasons for the crisis.

Exercise 1.3 (Bank run)
Explain what is meant by a bank run and give examples from history.

Exercise 1.4 (Regulatory frameworks)

a) Explain how the three-pillar concept works in the Basel and Solvency II regulatory frameworks.
b) Compare and contrast the main features of the Basel and Solvency II frameworks.

Exercise 1.5 (Procyclicality)
Explain the meaning of procyclicality, particularly when used in relation to regulatory capital requirements.

Exercise 1.6 (Basel III)
Summarize the main changes to the Basel framework that are adopted in Basel III and explain how these address shortcomings of the previous Basel II Accord.

Exercise 1.7 (Capital adequacy, leverage and liquidity coverage ratios)
Explain the essential differences between capital adequacy ratios, leverage ratios and liquidity coverage ratios as applied in the Basel III accord.

Case Studies

Exercise 1.8 (Major failures)
Select from the following list of well-known failures of financial risk management. In each case, summarize what happened and identify the risks that the company or organisation failed to manage.

a) Barings Bank
b) Metallgesellschaft (MG)
c) Sumitomo
d) Société Générale
e) Orange County
f) Long-Term Capital Management (LTCM)
g) American International Group (AIG)
h) Amaranth Advisors

Exercise 1.9 (Market crashes)
Investigate what happened during the following well-known market crashes and describe any changes that have resulted from these events, or any lessons that risk managers should draw.

a) The Wall Street Crash of 1929.
b) The 'Black Monday' event of 19 October 1987.
c) The 'Flash Crash' of 6 May 2010.

Exercise 1.10 (Natural catastrophes)
Research the following natural catastrophes and describe their economic, financial and insurance consequences.

a) The Kobe Earthquake (1995).
b) Hurricane Katrina (2005).
c) The Thailand Floods (2011).

Discussion

Exercise 1.11 (Risk and uncertainty)
Attempt to formulate your own succinct definition of risk. What distinguishes risk from uncertainty in your opinion?

Exercise 1.12 (The Q in QRM)
Discuss whether quantitative methodology should play a larger or a more reduced role in risk management systems in the future.

Exercise 1.13 (Trends in financial regulation)
What are the current trends in financial regulation? In particular:

a) Has risk regulation become too complex?
b) Is there evidence of a movement away from internal models back to simpler standardized approaches?

Exercise 1.14 (Regulation and credit provision)
An often-heard criticism of Basel III by banking practitioners is that "increased regulation strangles credit provision". Discuss the arguments for and against this proposition.

Exercise 1.15 (Shadow banking and insurance)
a) What is meant by the shadow banking industry?
b) Is there also a shadow insurance industry?

2 Basic Concepts in Risk Management

Review

Exercise 2.1 (Notions of capital)

In each of the following situations a certain notion of capital discussed in MFE (2015, Section 2.1.3) is most relevant for the decision maker. Explain which notion of capital that is.

a) A financial analyst who uses balance-sheet data to value a firm.

b) A chief risk officer of an insurance company who has to decide on the appropriate level of reinsurance.

c) A regulator who has to decide on shutting down a bank with many bad loans on its book.

Exercise 2.2 (Different notions of financial distress)

a) Briefly explain the difference between illiquidity, insolvency, default and bankruptcy.

b) Describe a scenario where a financial company is insolvent but has not defaulted.

c) Describe a scenario where a financial company has defaulted but is not insolvent.

Exercise 2.3 (Valuation of a real-estate investment)

Suppose the manager of a real-estate fund has to value a particular flat in Zurich, Switzerland. Relate the following three methods for the valuation of the flat to the valuation methods discussed in MFE (2015, Section 2.2.2).

a) The manager takes the purchase price of the flat from several years ago and reduces it by an annual depreciation of 1% to allow for wear and tear.

b) The manager finds transaction prices for similar flats in the neighborhood and uses them to compute a price per square metre. She then multiplies the square-metre price by the size of the flat.

c) The manager estimates prices per square meter from a broad Swiss property price index and makes an ad hoc adjustment of 20% to account for the location of the flat in Zurich.

Exercise 2.4 (Translation invariance of risk measures)

Show that a translation-invariant risk measure can be interpreted as the amount of capital that needs to be added to a position so that it becomes acceptable to a regulator.

Exercise 2.5 (Subadditivity of risk measures)

Explain why subadditivity is often considered a desirable property of a risk measure.

Exercise 2.6 (VaR and expected shortfall)

a) Give mathematically precise definitions of value-at-risk $\mathrm{VaR}_\alpha(L)$ and expected shortfall $\mathrm{ES}_\alpha(L)$ for a random loss L at confidence level $\alpha \in (0, 1)$.

b) Explain the relative advantages of each risk measure over the other.

Exercise 2.7 (Superadditivity scenarios for VaR)

Describe some models for financial losses that can lead to situations where VaR_α is superadditive.

Exercise 2.8 (Additivity for two linearly dependent random variables)

Consider an arbitrary random variable X and let $Y = aX + b$ for constants $a > 0$ and $b \in \mathbb{R}$. Show that $\mathrm{VaR}_\alpha(X + Y) = \mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y)$ for $\alpha \in (0, 1)$.

Basic

Exercise 2.9 (Risk-neutral valuation for interest-rate derivatives)

Consider a two-period model. Denote by r_t, $t \in \{0, 1\}$, the simple interest rate from t to $t + 1$, so that 1 monetary unit invested at t is worth $1 + r_t$ at $t + 1$. Assume that r_0 is 1.5% and that r_1 takes the values 1% and 2% with probability 1/2. Denote by $p(t, T)$ the price at t of a zero-coupon bond with maturity T and face value 1.

a) Write down $p(0, 1)$ and $p(1, 2)$ for the cases $r_1 = 0.01$ and $r_1 = 0.02$.

b) Suppose a long zero-coupon bond with maturity $T = 2$ and face value 1 is traded for 0.969729 at $t = 0$. In this setup an equivalent martingale measure $\mathbb{Q}$ is characterized by the probability $q = \mathbb{Q}(r_1 = 0.01)$. Compute q from $p(0, 2)$.

c) Apply risk-neutral valuation to price a stylized floor contract which pays an amount of 1 if $r_1 < r_0$.

   Note. In general, a *floor contract* is an option which provides protection against low interest rates.

Exercise 2.10 (Mapping of a stock portfolio affected by exchange rates)

Consider a portfolio $\mathcal{P}$ consisting of two stocks $S_{t,1}, S_{t,2}$, where $S_{t,1}$ denotes the value of stock 1 in EUR and $S_{t,2}$ denotes the value of stock 2 in CHF. Let e_t^{CHF} denote the CHF/EUR exchange rate at time t. In other words, 1 CHF is worth e_t^{CHF} EUR at t. Furthermore, denote by λ_1 and λ_2 the number of shares in stocks 1 and 2 in $\mathcal{P}$, respectively.

a) Derive the value V_t in EUR of $\mathcal{P}$ at time t in terms of the risk factors $Z_{t,j} = \log S_{t,j}$, $j \in \{1, 2\}$, and $Z_{t,3} = \log e_t^{\mathrm{CHF}}$. What is the corresponding mapping?

b) Derive the value V_{t+1} of $\mathcal{P}$ at time $t + 1$ and the one-period loss L_{t+1}.

c) Derive the linearized one-period loss L_{t+1}^{Δ} and express it in terms of portfolio weights w_1, w_2 (the values of each stock investment relative to the value V_t of the overall portfolio).

Exercise 2.11 (Properties of VaR and expected shortfall)

a) Show that VaR_α is a monotone, translation invariant and positive-homogeneous risk measure.

b) Why can we conclude that ES_α also satisfies these properties?

Exercise 2.12 (VaR and ES for continuous distributions with finite mean)
Compute $\mathrm{VaR}_\alpha(L)$ and $\mathrm{ES}_\alpha(L)$ in the following cases:

a) L has an exponential distribution with rate parameter $\lambda > 0$, that is $L \sim \mathrm{Exp}(\lambda)$.

b) L has a log-normal distribution, that is $\log L \sim \mathrm{N}(\mu, \sigma^2)$.

c) L has the Weibull distribution with distribution function $F(x) = 1 - \exp(-\sqrt{x}/10)$, $x \geq 0$.

d) L has a Pareto distribution $\mathrm{Pa}(\theta, 1)$ with distribution function $F(x) = 1 - (1 + x)^{-\theta}$, $x \geq 0$, $\theta > 0$. Under what condition on θ does $\mathrm{ES}_\alpha(L)$ exist?

Exercise 2.13 (VaR and ES for a distribution function with jumps)
Suppose that the loss L has distribution function

$$F(x) = \begin{cases} 0, & x < 1, \\ 1 - 1/(1 + x), & x \in [1, 3), \\ 1 - 1/x^2, & x \geq 3. \end{cases} \tag{E1}$$

a) Plot the graph of F.

b) Compute value-at-risk $\mathrm{VaR}_\alpha(L)$ at confidence levels $\alpha = 85\%$ and $\alpha = 95\%$.

c) Compute expected shortfall $\mathrm{ES}_\alpha(L)$ at confidence level $\alpha = 85\%$.

Exercise 2.14 (VaR for a binomial model of a stock price)
Consider a portfolio consisting of a single stock with current value $S_t = 100$. Each year, the stock price either increases by 4% with probability 0.8 or decreases by 4% with probability 0.2. Compute VaR_α for $\alpha \in \{0.7, 0.95, 0.96, 0.99\}$ over a time horizon of two years.

Exercise 2.15 (VaR and ES for a discrete distribution)
The following table contains the net profits on two lines of business A and B of a company XYZ:

		Net profit	
Outcome	Probability	Line A	Line B
ω_1	0.82	1000	1000
ω_2	0.04	1000	0
ω_3	0.04	0	1000
ω_4	0.02	0	0
ω_5	0.04	1000	−10 000
ω_6	0.04	−10 000	1000

a) Calculate $\mathrm{VaR}_{0.95}$ and $\mathrm{ES}_{0.95}$ for each of the business lines A and B.

b) Calculate $\mathrm{VaR}_{0.95}$ and $\mathrm{ES}_{0.95}$ for the combined profits of A and B.

How do your answers fit in with your knowledge of the coherence of the value-at-risk and expected shortfall risk measures?

Exercise 2.16 (Expected shortfall for *t* distributions)
Compute ES_α for a standard t distribution with ν degrees of freedom, $L \sim t_\nu$. Give a condition that guarantees that $\mathrm{ES}_\alpha(L)$ exists.

Exercise 2.17 (VaR and ES for bivariate normal risks)
Consider two stocks whose log-returns are bivariate normally distributed with annualized volatilities $\sigma_1 = 0.2$, $\sigma_2 = 0.25$ and correlation $\rho = 0.4$. Assume that the expected returns are equal to 0 and that one year consists of 250 trading days. Consider a portfolio with current value $V_t = 10^6$ (in EUR) and portfolio weights $w_1 = 0.7$ and $w_2 = 0.3$. Furthermore, denote by L^Δ_{t+1} the linearized loss (using log-prices of the stocks as risk factors). Compute the daily $\mathrm{VaR}_{0.99}(L^\Delta_{t+1})$ and $\mathrm{ES}_{0.99}(L^\Delta_{t+1})$ for the portfolio. How does the answer change if $\rho = 0.6$?

Exercise 2.18 (Basic historical simulation)
Consider two stocks, indexed by $j \in \{1, 2\}$, with current values $S_{t,1} = 1000$ and $S_{t,2} = 100$ in some unit of currency. The monthly log-returns of both stocks over the last 10 months are given in the following table (values in %):

Lag k	10	9	8	7	6	5	4	3	2	1
Log-return of S_1 at lag k	−16.1	5.1	−0.4	−2.5	−4	10.5	5.2	−2.9	19.1	0.4
Log-return of S_2 at lag k	−8.2	3.1	0.4	−1.5	−3	4.5	2	−3.7	10.9	−0.4

Table E.2.1 Monthly log-returns of the two stocks over the last 10 months (in %).

Use historical simulation to estimate the one-month VaR at confidence level $\alpha = 0.9$ for the linearized loss L^Δ for the following two portfolios:

a) A portfolio consisting of two shares of the first stock S_1.
b) A portfolio consisting of one share of the first stock and 10 shares of the second stock.

Exercise 2.19 (Axioms of coherence for standard deviation)
Consider the risk measure $\varrho(L) = \sqrt{\operatorname{var} L}$ on the space $\mathcal{M} = L^2(\Omega, \mathcal{F}, \mathbb{P})$ of random variables with finite second moment. For each axiom of coherence, either prove it or provide a counterexample to show it does not hold.

Exercise 2.20 (Superadditivity of VaR for two iid Bernoulli random variables)
Let L_1, L_2 be independent and identically distributed (iid) Bernoulli random variables such that $\mathbb{P}(L_j = 0) = 1 - p$ and $\mathbb{P}(L_j = 1) = p$, $j \in \{1, 2\}$, for some $p \in (0, 1)$. For which $\alpha \in (0, 1)$ is VaR_α superadditive, that is $\mathrm{VaR}_\alpha(L_1 + L_2) > \mathrm{VaR}_\alpha(L_1) + \mathrm{VaR}_\alpha(L_2)$?

Exercise 2.21 (VaR under strictly increasing transformations of a loss)
Let L be a random loss and $h : \mathbb{R} \to \mathbb{R}$ a continuous and strictly increasing function. Show that

$$\mathrm{VaR}_\alpha(h(L)) = h(\mathrm{VaR}_\alpha(L)), \quad \alpha \in (0, 1).$$

Exercise 2.22 (Subadditivity of VaR for bivariate normal random variables)
Let (L_1, L_2) have a bivariate normal distribution with $\mu_j = \mathbb{E}(L_j)$, $\sigma_j^2 = \mathrm{var}(L_j)$, $j \in \{1, 2\}$, and $\rho = \mathrm{corr}(L_1, L_2)$.

a) Compute $\mathrm{VaR}_\alpha(L_j)$, $j \in \{1, 2\}$, and $\mathrm{VaR}_\alpha(L_1 + L_2)$.
b) Show that VaR_α is subadditive if and only if $\alpha \in [1/2, 1)$.
c) Plot the graphs of $\alpha \mapsto \mathrm{VaR}_\alpha(L_1 + L_2)$ for standard normal L_1, L_2 with correlation values $\rho \in \{-1, -0.5, 0, 0.5, 1\}$. Comment on the monotonicity behaviour with respect to ρ.

Exercise 2.23 (Shortfall-to-quantile ratio for exponential, normal and Pareto)
Calculate the ratio

$$\lim_{\alpha \to 1-} \frac{\mathrm{ES}_\alpha(L)}{\mathrm{VaR}_\alpha(L)}$$

in the following cases. Comment on the differences from a risk management perspective.

a) $L \sim \mathrm{Exp}(\lambda)$ with distribution function $F(x) = 1 - \exp(-\lambda x)$, $x \geq 0$, $\lambda > 0$ (a light-tailed distribution).
b) $L \sim \mathrm{N}(0, 1)$ (also light-tailed).
c) $L \sim \mathrm{Pa}(\theta, 1)$ with distribution function $F(x) = 1 - (1 + x)^{-\theta}$, $x \geq 0$, $\theta > 1$ (a heavy-tailed distribution).

Exercise 2.24 (Matching VaR and ES)
Under the Basel III guidelines, the risk measure underlying the capital requirements for market risk has changed from VaR to expected shortfall; see BIS (2016) and BIS (2017). In order that the overall capital requirement remains constant, the confidence level of ES can be adjusted.

Use numerical root finding to determine the value of α such that $\mathrm{ES}_\alpha(L) = \mathrm{VaR}_{0.99}(L)$ in the following cases.

a) $L \sim \mathrm{N}(\mu, \sigma^2)$.

   Hint. First show that it is sufficient to consider $L \sim \mathrm{N}(0, 1)$.
b) $L \sim t_{3.5}$.

Exercise 2.25 (Matching VaR and ES for Pareto distributions)
Let $\beta \in (0, 1)$ and $L \sim \mathrm{Pa}(\theta, 1)$ with distribution function $F(x) = 1 - (1 + x)^{-\theta}$, $x \geq 0$, $\theta > 1$.

a) Determine analytically the confidence level $\alpha \in (0, \beta]$ such that $\mathrm{ES}_\alpha(L) = \mathrm{VaR}_\beta(L)$ and give the range of β values for which the equation has a solution.
b) Suppose that β is such that the equation has a solution. Show that α is given by an increasing function of θ and explain why this is the case.

Advanced

Exercise 2.26 (Superadditivity of VaR for two iid exponential random variables)
For $j \in \{1,2\}$ let F_j denote the distribution function of L_j and let $F_{L_1+L_2}$ denote the distribution function of $L_1 + L_2$.

a) Show that superadditivity of VaR_α is equivalent to

$$\alpha > F_{L_1+L_2}(F_{L_1}^{\leftarrow}(\alpha) + F_{L_2}^{\leftarrow}(\alpha)).$$

b) Now assume that L_1 and L_2 are iid exponentially distributed with distribution function $F(x) = 1 - \exp(-\lambda x)$, $x \geq 0$, $\lambda > 0$. Show that VaR_α is superadditive if and only if $(1-\alpha)(1 - 2\log(1-\alpha)) > 1$.

c) Determine the set of α values for which VaR_α is superadditive using numerical root finding.

Exercise 2.27 (Superadditivity of VaR for iid Bernoulli random variables)
For $d \geq 2$ let $Y_1, \ldots, Y_d$ be iid Bernoulli risks such that $Y_j \sim \mathrm{B}(1,p)$ for $p \in [0,1]$. Show that VaR_α is superadditive if and only if $(1-p)^d < \alpha \leq 1 - p$.

Exercise 2.28 (Superadditivity of VaR for a heavy tailed distribution)
Let L_1 and L_2 be independent random variables with distribution function $F(x) = 1 - x^{-1/2}$, $x \in [1, \infty)$. This is a form of Pareto distribution with infinite mean. Show that $\mathrm{VaR}_\alpha(L_1 + L_2) > \mathrm{VaR}_\alpha(L_1) + \mathrm{VaR}_\alpha(L_2)$ for all $\alpha \in (0,1)$.

Hint. Show that the distribution function $F_{L_1+L_2}$ of $L_1 + L_2$ is given by $F_{L_1+L_2}(x) = 1 - 2\sqrt{x-1}/x$, $x \geq 2$.

Exercise 2.29 (Median shortfall)
Let the loss L have continuous distribution function F_L. The risk measure *median shortfall* at confidence level $\alpha \in [0,1)$ is given by

$$\mathrm{MS}_\alpha(L) = F^{\leftarrow}_{L|L>\mathrm{VaR}_\alpha(L)}(1/2),$$

where $F_{L|L>\mathrm{VaR}_\alpha(L)}$ denotes the conditional distribution of L given that L exceeds $\mathrm{VaR}_\alpha(L)$.

a) Express $F_{L|L>\mathrm{VaR}_\alpha(L)}$ in terms of F_L.

b) Show that $\mathrm{MS}_\alpha(L) = \mathrm{VaR}_{\frac{1+\alpha}{2}}(L)$ and interpret this result.

c) Let $F_L(x) = 1 - (1+x)^{-\theta}$, $x \geq 0$, $\theta > 1$, be the distribution function of the Pareto distribution. Compute

$$\lim_{\alpha \to 1-} \frac{\mathrm{ES}_\alpha(L)}{\mathrm{MS}_\alpha(L)}$$

and interpret this result from a risk management perspective.

Exercise 2.30 (Shortfall-to-quantile ratio for t)
Let $L \sim t_\nu(\mu, \sigma^2)$ for $\nu > 1$, $\mu \in \mathbb{R}$, $\sigma > 0$. Compute $\mathrm{VaR}_\alpha(L)$ and $\mathrm{ES}_\alpha(L)$. Show that

$$\lim_{\alpha \to 1-} \frac{\mathrm{ES}_\alpha(L)}{\mathrm{VaR}_\alpha(L)} = \frac{\nu}{\nu - 1}$$

and interpret this result from a risk management perspective.

3 Empirical Properties of Financial Data

Review

Exercise 3.1 (Stylized facts)

Give a list of univariate stylized facts for financial time series.

Exercise 3.2 (Negative log-returns and the correlogram)

Figure E.3.1 shows the negative log-returns of the S&P 500 index from 3 January 2011 to 31 December 2015 (left) and the corresponding correlogram or sample autocorrelation function (ACF) plot (right).

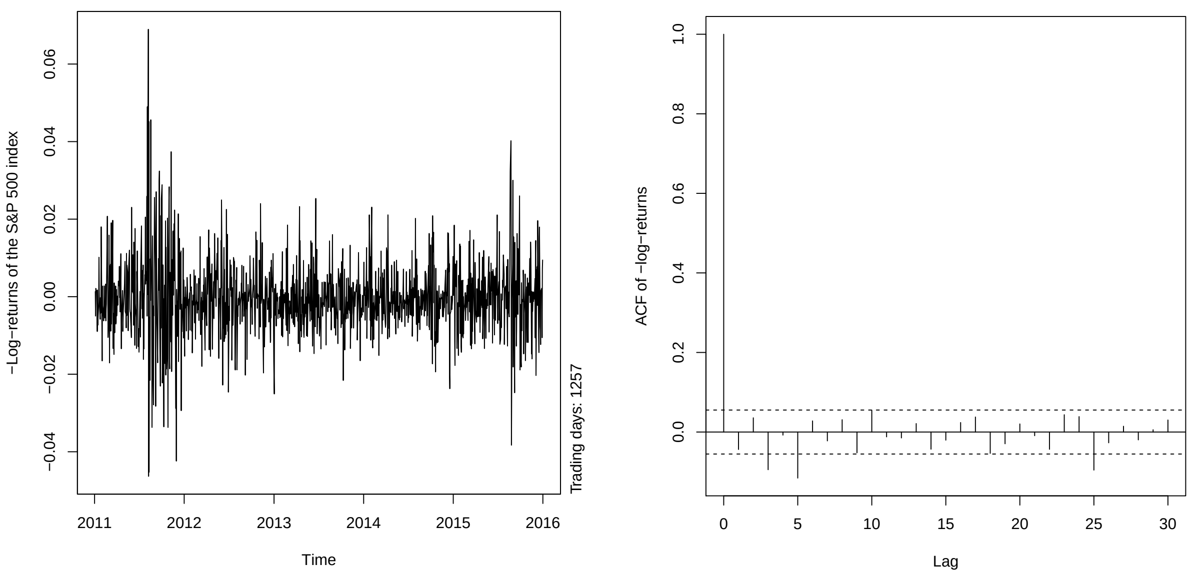

Figure E.3.1 Negative log-returns of the S&P 500 index from 3 January 2011 to 31 December 2015 (left) and the corresponding correlogram (right).

a) Explain how the correlogram in Figure E.3.1 should be interpreted.
b) Which stylized fact(s) do the two plots in Figure E.3.1 support?
c) Explain why it is not sufficient to test whether a dataset satisfies the iid hypothesis by looking at the correlogram of the dataset alone. What else should you do?

Exercise 3.3 (Interpreting Q-Q plots)

Figure E.3.2 shows a Q-Q plot of the negative log-returns of the price of one Bitcoin in USD from 29 May 2014 to 29 May 2018 against a fitted normal distribution.

a) What does this Q-Q plot suggest about the distribution of Bitcoin log-returns?

b) What numerical test might you carry out to support your answer to part a)?

c) Although the QQ-plot is against a fitted normal distribution, the ordered data could also have been plotted against the quantiles of a standard normal. Explain why this is the case.

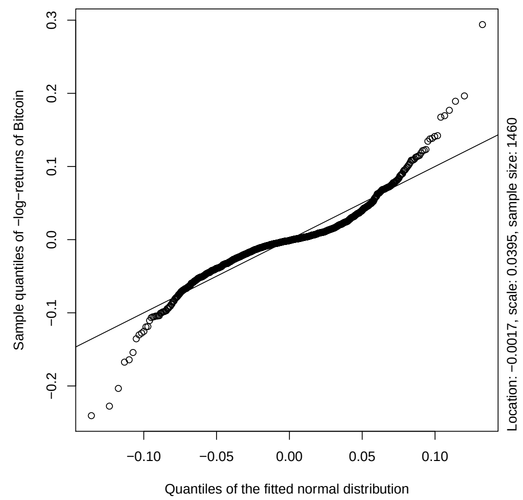

Figure E.3.2 Q-Q plot of Bitcoin returns against a normal distribution.

Exercise 3.4 (Multivariate stylized facts)

Give a list of multivariate stylized facts for financial time series.

Exercise 3.5 (Joint negative log-returns and the cross-correlogram)

Figure E.3.3 shows joint negative log-returns of the S&P 500 and the DAX index from 3 January to 31 December 2015 (left) and the corresponding cross-correlogram (right).

a) Based on your knowledge of the correlogram, how do you think the cross-correlogram in Figure E.3.3 should be interpreted?

b) Which stylized fact(s) do the two plots in Figure E.3.3 support?

c) How would you expect the cross-correlogram of the squared negative log-returns to look?

Exercise 3.6 (Longer-interval log-returns)

Let $S_t, \dots, S_{t+h}$ denote the daily values of an asset over a period of $h > 1$ days and $X_{t+1}, \dots, X_{t+h}$ the corresponding daily log-returns.

a) Give a formula for the h-day log-return $X_{t+h}^{(h)}$ in terms of $X_{t+1}, \dots, X_{t+h}$ for $h \geq 1$.

b) Explain why $X_{t+h}^{(h)}$ may be approximately normally distributed for large h (corresponding to a quarter or year) even if the daily returns are not.

Basic

Exercise 3.7 (Manipulating logarithmic and relative returns)

For an asset price with daily values given by the time series (S_t) the h-day log-return is defined by $X_{t+h}^{(h)} = \log(S_{t+h}/S_t)$ and the h-day relative return by $Y_{t+h}^{(h)} = (S_{t+h} - S_t)/S_t$ for all $h \geq 1$.

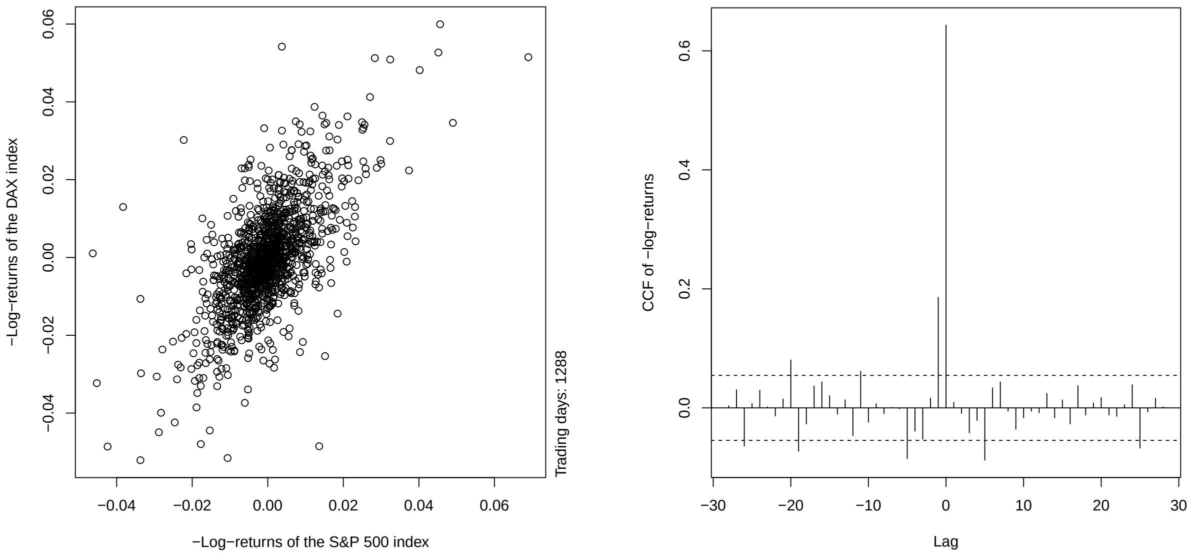

Figure E.3.3 Joint negative log-returns of the S&P 500 and DAX index from 3 January 2011 to 31 December 2015 (left) and the corresponding cross-correlogram (right).

a) Derive a formula for the h-day log-return $X^{(h)}_{t+h}$ in terms of the h-day relative return $Y^{(h)}_{t+h}$.

b) Derive a formula for the h-day relative return $Y^{(h)}_{t+h}$ in terms of the h-day log-return $X^{(h)}_{t+h}$.

c) Given a series of h-day log-returns $X^{(h)}_{t+h}, X^{(h)}_{t+2h}, \dots$ and an initial stock price S_t, explain how you would reconstruct the series of prices $S_{t+h}, S_{t+2h}, \dots$ at h-day intervals.

d) Given a series of h-day relative returns $Y^{(h)}_{t+h}, Y^{(h)}_{t+2h}, \dots$ and an initial stock price S_t, explain how you would reconstruct the series of prices $S_{t+h}, S_{t+2h}, \dots$ at h-day intervals.

Exercise 3.8 (Experimenting with Q-Q plots)
Use statistical software to generate 1000 data points from the following distributions; the quantity of data corresponds approximately to four years of daily return data. In each case form Q-Q plots against the standard normal distribution and interpret the results.

a) A Student t distribution with 6 degrees of freedom.

b) A standard uniform distribution.

c) A standard logistic distribution (with distribution function $F(x) = 1/(1+\exp(-x))$, $x \in \mathbb{R}$).

d) A standard lognormal distribution (that is the distribution of a random variable whose logarithm follows a standard normal distribution).

Exercise 3.9 (Ljung-Box test)
Repeat the analysis of MFE (2015, Table 3.1) for the stocks that comprise the Euro Stoxx 50 index. Specifically you should:

a) Form daily and monthly log-returns for each time series from 2000 to 2015, omitting all stocks with missing data on the first trading day in 2000 and filling the remaining missing data by interpolation.

b) Carry out Ljung-Box tests on both the raw and absolute returns to address the null hypothesis that stock returns are iid.

The data can be found in the dataset `EURSTX_const` of the R package `qrmdata`. The function `na.fill()` of the R package `zoo` might be helpful.

Exercise 3.10 (Jarque-Bera test)

Repeat the analysis of MFE (2015, Table 3.2) for the stocks that comprise the Euro Stoxx 50 index using the same data as in Exercise 3.9. Specifically you should:

a) Form weekly, monthly and quarterly log-returns.

b) Carry out Jarque-Bera tests of normality for each stock and each time step.

Exercise 3.11 (Skewness, kurtosis and higher moments of standard normal)

a) Let Z be a random variable with a density f that is symmetric around 0. Show that if the kth moment exists it is given by

$$\mathbb{E}(Z^k) = \begin{cases} 2\int_0^\infty z^k f(z)\,\mathrm{d}z, & k \text{ is even}, \\ 0, & k \text{ is odd}. \end{cases}$$

b) Let $Z \sim \mathrm{N}(0,1)$. By writing $\int_0^\infty z^k \phi(z)\,\mathrm{d}z$ in terms of the gamma function, derive the general formula

$$\mathbb{E}(Z^k) = \begin{cases} \prod_{j=1}^{k/2}(k-(2j-1)), & k \text{ is even}, \\ 0, & k \text{ is odd}. \end{cases}$$

Hint. You may use the fact that $\Gamma(x+1) = x\Gamma(x)$, $x > 0$, for the gamma function and the identity $\Gamma(1/2) = \sqrt{\pi}$.

c) Let X be a random variable with a finite kth moment. The kth standardized moment of X is defined by $\mathbb{E}((X-\mu)^k)/\sigma^k$, where $\mu = \mathbb{E}(X)$ and $\sigma = \sqrt{\mathrm{var}(X)}$ denote the mean and standard deviation of X. The third and fourth standardized moments of X are known as the skewness β and the kurtosis κ. Compute the skewness, kurtosis and the 6th standardized moment of a normal random variable $X \sim \mathrm{N}(\mu, \sigma^2)$.

Exercise 3.12 (Changing correlation over time)

Repeat the analysis of MFE (2015, Figure 3.7) for the S&P 500 and DAX log-returns from 3 January 2011 to 31 December 2015 (the data used in Exercise 3.5). Specifically you should:

a) Estimate correlations using a rolling 25-day window and plot the time series of estimated correlations.

b) Experiment with different window sizes and investigate how the plot changes.

The data can be found in the datasets `SP500` and `DAX` of the R package `qrmdata`.

Advanced

Exercise 3.13 (A simple version of the runs test)

In this question we study a simple version of a test for iid data known as the *runs test*. Let $X_1, \dots, X_n$ be return data from a continuous distribution that is approximately symmetric around zero. Let $Y_t = \operatorname{sign}(X_t)$ denote the sign of X_t. Throughout this question we will assume that $\mathbb{P}(Y_t = -1) = \mathbb{P}(Y_t = 1) = 1/2$.

The runs test is based on the idea of counting the number of runs, R_n, of $+$ values and $-$ values in $Y_1, \dots, Y_n$. For example if the whole sequence is $+ + - - - + - - + + +$ then the number of runs is 5.

a) Explain why the number R_n of runs is given by

$$R_n = 1 + \sum_{t=2}^{n} \frac{1}{4}(Y_t - Y_{t-1})^2.$$

b) Show that $\sum_{t=2}^{n} \frac{1}{4}(Y_t - Y_{t-1})^2$ has a binomial distribution under the null hypothesis that $X_1, \dots, X_n$ are iid. Determine the parameters of the binomial distribution.

c) Using the normal approximation to the binomial distribution, explain how you could implement a test of the null hypothesis of independence based on the statistics R_n.

d) Suppose we find no evidence against the iid hypothesis for the return data $X_1, \dots, X_n$. How could the runs test be adapted to test for clustering of large values in the squared data?

Note. The standard Wald–Wolfowitz runs test does not assume that $\mathbb{P}(Y_t = -1) = \mathbb{P}(Y_t = 1) = 1/2$ and uses a slightly different procedure to the one described above; see, for example, Bradley (1968, Chapter VIII).

Exercise 3.14 (Sensitivity of correlation estimates to volatility)

This exercise is a continuation of Exercise 3.12. Repeat the analysis of MFE (2015, Figure 3.8) for the S&P 500 and DAX log-returns from 3 January 2011 to 31 December 2015. Specifically you should:

a) Carry out a regression analysis to investigate whether correlation estimates calculated for non-overlapping 25-day blocks of observations tend to be higher in periods with higher volatility estimates.

b) Repeat the analysis using rolling estimates of correlation and volatility.

Exercise 3.15 (Geometric spacings between largest values in iid samples)

This exercise explains the theory underlying MFE (2015, Figure 3.3). Let $X_1, X_2, \dots$ be a sequence of iid random variables from a continuous distribution function F and that $p = \mathbb{P}(X_1 > u) = \bar{F}(u) \in (0,1)$ for some threshold u. Let $T_0 = 0$ and $T_j = \min\{i > T_{j-1} : X_i > u\}$, $j \in \mathbb{N}$.

a) Show that the spacings $S_j = T_j - T_{j-1}$, $j \in \mathbb{N}$, of exceedances of the X_i's over the threshold u form a series of iid random variables following a geometric distribution with parameter p.

b) Explain how the insight in a) can be used to implement a graphical test for volatility clustering?

c) Show that the geometric distribution can be approximated by an exponential distribution with parameter p if p is small.

d) Use computer software to investigate the quality of the exponential approximation for $p \in \{0.01, 0.02, 0.04\}$.

Exercise 3.16 (Log-returns implied by geometric Brownian motion)

The geometric Brownian motion is a widely used model for stock price processes in mathematical finance. According to this model, the stock price process $(S_t)_{t\geq 0}$ satisfies $S_t = S_0 \exp((\mu - \sigma^2/2)t + \sigma W_t)$ for $\mu \in \mathbb{R}$, $\sigma^2 > 0$, where the stochastic process $(W_t)_{t\geq 0}$ satisfies:

1) $W_0 = 0$ almost surely and $t \mapsto W_t$ is continuous almost surely;
2) For any $n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \cdots < t_n$, the increments $W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent; and
3) For all $0 \leq s < t$, $W_t - W_s \sim \mathrm{N}(0, t-s)$.

Answer the following questions:

a) What model does this assumption imply for the distribution of the h-period risk-factor change $X_{t+h}^{(h)} = -\log(S_{t+h}/S_t)$? Is this model supported by stylized facts about financial time series?

b) Show that the two risk-factor changes $X_{t+h}^{(h)}$ and $X_{(t+h)+h'}^{(h')}$ for $h, h' \geq 0$ are independent and comment on whether this fact is supported by empirical properties of financial time series?

4 Financial Time Series

Review

Exercise 4.1 (Stationarity)

a) State the definitions of covariance (or weak) stationarity and strict stationarity for a time series $(X_t)_{t\in\mathbb{Z}}$.

b) Give an example of a covariance-stationary process that is not strictly stationary and vice versa.

c) Explain why stationarity is such an important assumption in time series analysis.

d) If you had time series data displaying a linear trend, explain how you might pre-process them to make them suitable for modelling with a stationary time series process.

Exercise 4.2 (Identification of ARMA processes)

Figure E.4.1 shows simulated paths for various time series processes as well as the corresponding sample autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. In each case the innovations $(\varepsilon_t)_{t\in\mathbb{Z}}$ are from a strict white noise process and the time series $(X_t)_{t\in\mathbb{Z}}$ follows one of the following models:

a) $X_t = \varepsilon_t - 0.3\varepsilon_{t-1} - 0.4\varepsilon_{t-2}$, $t \in \mathbb{Z}$

b) $X_t = 10 + 20t + 2t\sin(t) + \varepsilon_t$, $t \in \mathbb{Z}$

c) $X_t - 0.3X_{t-1} - 0.4X_{t-2} = \varepsilon_t$, $t \in \mathbb{Z}$

d) $X_t = \varepsilon_t - 0.7\varepsilon_{t-1}$, $t \in \mathbb{Z}$

e) $X_t - 0.3X_{t-1} = \varepsilon_t - 0.6\varepsilon_{t-1}$, $t \in \mathbb{Z}$

Match each model with the pictures, explaining your reasoning.

Note. The PACF is mentioned briefly in MFE (2015, p. 108). The PACF of an AR(p) process cuts off after lag p and decays exponentially for an MA process. Note that this is the opposite behaviour to that of the ACF, which cuts off after lag q for an MA(q) process and decays exponentially for an AR process. Neither the ACF nor the PACF give clear guidance for indentifying the order of an ARMA(p,q) process with $p > 0$ and $q > 0$. See, for example, Brockwell and Davis (2002).

Exercise 4.3 (GARCH models and stylized facts)

Consider a GARCH(1, 1) process with parameters $\alpha_0 = 2$, $\alpha_1 = 0.1$ and $\beta_1 = 0.85$.

a) Discuss whether or not this process can replicate all of the univariate stylized facts of financial time series listed in MFE (2015, Section 3.1).

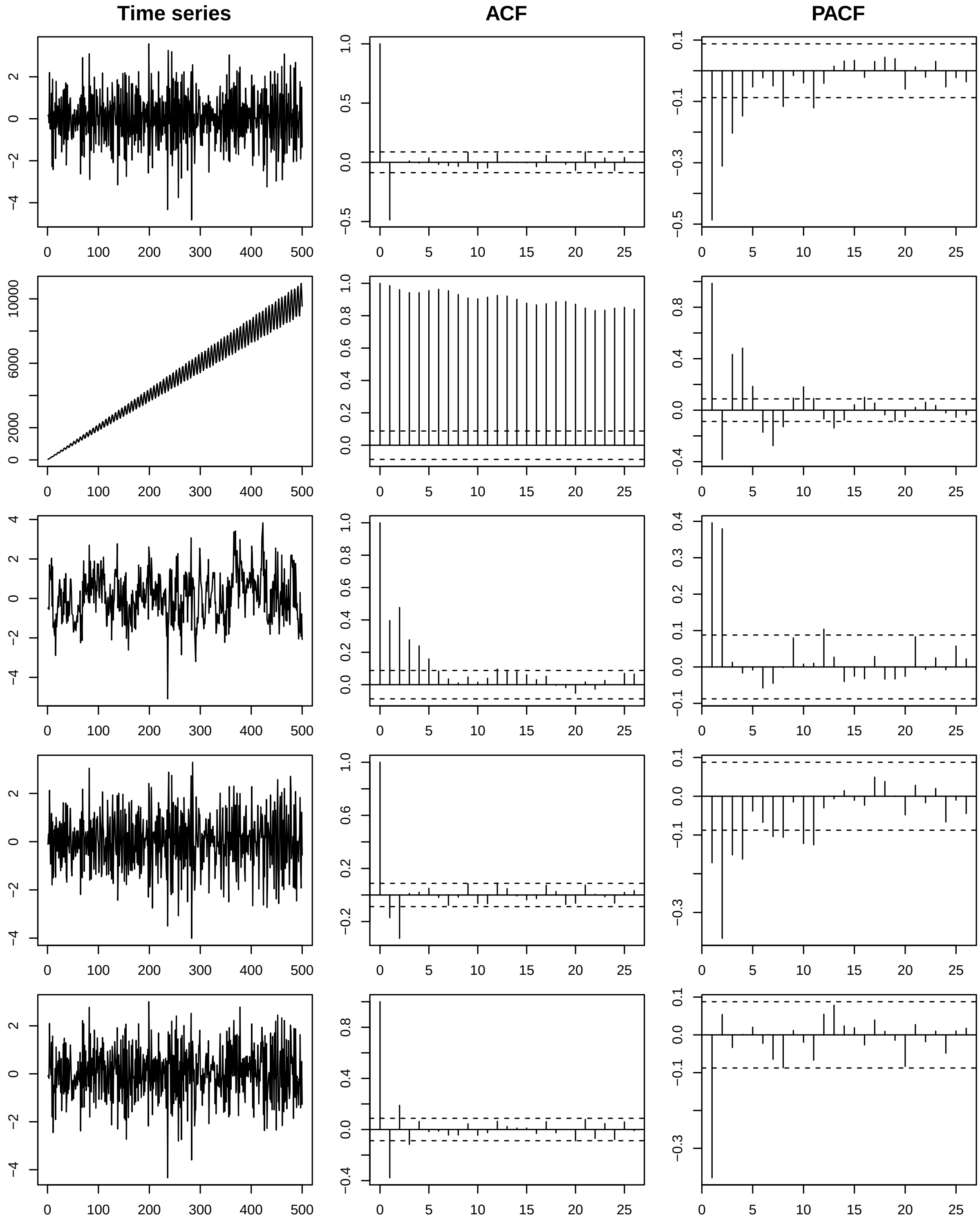

Figure E.4.1 Paths of five time series $(X_t)_{t\in\mathbb{Z}}$ (left) with corresponding ACF (middle) and PACF (right).

b) Calculate the variance of the GARCH process.

c) Explain the form that the autocorrelation function of the squared GARCH process would take.

Exercise 4.4 (White noise)

a) Explain the difference between a white noise process (WN) and a strict white noise process (SWN).

b) Give an example of a WN that is not SWN and a SWN that is not WN.

Exercise 4.5 (AR models with ARCH errors)

a) Write down the full set of equations for a time series $(X_t)_{t\in\mathbb{Z}}$ following an AR(1) model with non-zero mean and ARCH(1) errors.

b) Give a formula for value-at-risk calculated at time t, that is for the conditional quantile of X_{t+1} in terms of previous values of the process and quantiles of the innovation distribution.

Exercise 4.6 (QML estimation of GARCH models)
Explain how the quasi-maximum likelihood (QML) method for estimating the parameters of a GARCH process works and contrast the properties of the resulting QML estimates with those of full maximum likelihood estimates.

Basic

Exercise 4.7 (Checking stationarity)
Decide whether the following time series (X_t) are covariance stationary. Calculate the mean function and autocorrelation function of any time series that is covariance stationary.

a) $X_t = \begin{cases} \varepsilon_t & \text{if } t \text{ is odd,} \\ \varepsilon_{t+1} & \text{if } t \text{ is even,} \end{cases}$ where $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a $\mathrm{WN}(0,\sigma_\varepsilon^2)$ process.

b) $X_t = \sum_{k=0}^{t} \varepsilon_k$ for $t \in \mathbb{N}_0$, where $(\varepsilon_t)_{t\in\mathbb{N}_0}$ is a $\mathrm{WN}(0,\sigma_\varepsilon^2)$ process.

c) $X_t = \sin\big(\frac{\pi}{6}t\big)Z_2 + \cos\big(\frac{\pi}{6}t\big)Z_1$, $t \in \mathbb{Z}$, where Z_1 and Z_2 are independent random variables with mean 0 and variance σ_Z^2.

Exercise 4.8 (Symmetry of ACF)
Show that the autocorrelation function ρ of a covariance-stationary time series is an even function, that is $\rho(-h) = \rho(h)$, $h \in \mathbb{Z}$.

Exercise 4.9 (MA(2) process)
Let $(X_t)_{t\in\mathbb{Z}}$ be a time series of the form

$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}, \quad t \in \mathbb{Z},$$

where $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,\sigma_\varepsilon^2)$.

a) Show from first principles (that is without using general formulas provided in MFE (2015)) that $(X_t)_{t\in\mathbb{Z}}$ is causal and stationary and compute the mean function and autocorrelation function.

b) Assume that $(X_t)_{t\in\mathbb{Z}}$ is an invertible process. Derive formulas for the predictions $P_t X_{t+h}$ for $h = 1$ and $h = 2$ and explain what happens when $h \geq 3$.

Exercise 4.10 (Absolute summability condition)

Let $X_t = \sum_{k=-\infty}^{\infty} \psi_k \varepsilon_{t-k}$ for $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0, \sigma_\varepsilon^2)$. Show that $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$ implies that $\mathbb{E}(\sum_{k=-\infty}^{\infty} |\psi_k \varepsilon_{t-k}|) < \infty$ or, in other words, $\sum_{k=-\infty}^{\infty} \psi_k \varepsilon_{t-k}$ converges absolutely, almost surely.

Exercise 4.11 (AR(1) process)

This question is intended to fill in some of the details and extend the analysis of the AR(1) model in MFE (2015, Example 4.11). Let the process $(X_t)_{t\in\mathbb{Z}}$ be a stationary solution of the AR(1) equations

$$X_t - \phi_1 X_{t-1} = \varepsilon_t, \quad t \in \mathbb{Z},\ |\phi_1| < 1, \tag{E1}$$

where $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0, \sigma_\varepsilon^2)$.

a) Show that if $(Y_t)_{t\in\mathbb{Z}}$ is another stationary solution of (E1), then the process $(Z_t)_{t\in\mathbb{Z}}$ defined by $Z_t = X_t - Y_t$ is almost surely equal to zero for all t.

   Hint. Derive the recurrence relation $Z_t = \phi_1^n Z_{t-n}$, $n \in \mathbb{N}$, and use the Cauchy–Schwarz inequality $\mathbb{E}(|Z_{t-n}|) \leq \sqrt{\mathbb{E}(Z_{t-n}^2)}$ to show that $\mathbb{E}(|Z_t|) = 0$ for all $t \in \mathbb{Z}$.

b) Let $Y_t = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-k}$, $t \in \mathbb{Z}$. Show that this series converges absolutely almost surely.

   Hint. Apply Exercise 4.10.

c) Show that $(Y_t)_{t\in\mathbb{Z}}$ satisfies (E1) and hence conclude from a) that $X_t = Y_t$, almost surely, for $t \in \mathbb{Z}$.

d) Compute the mean function and autocorrelation function of $(X_t)_{t\in\mathbb{Z}}$.

e) Derive a prediction formula for $P_t X_{t+h}$ for all $h \in \mathbb{N}$.

Exercise 4.12 (ARMA(1,1) process)

This question is intended to fill in some of the details of the analysis of the ARMA(1, 1) model in MFE (2015, Example 4.12). Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary and causal solution of

$$X_t - \phi_1 X_{t-1} = \varepsilon_t + \theta_1 \varepsilon_{t-1}, \quad t \in \mathbb{Z}, \tag{E1}$$

for some constants $\phi_1 \neq 0$, $\theta_1 \neq 0$ and $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0, \sigma_\varepsilon^2)$, $0 < \sigma_\varepsilon < \infty$.

a) Assume that $\phi_1 \neq -\theta_1$. By using the general ARMA theory in MFE (2015, p. 103) show that a necessary and sufficient condition for the equations (E1) to have a stationary and causal solution is $|\phi_1| < 1$.

b) By using MFE (2015, Proposition 4.9), compute the autocorrelation function of $(X_t)_{t\in\mathbb{Z}}$ in this case.

c) What is the solution of the equations (E1) when $\phi_1 = -\theta_1$?

Hint. One can proceed as in the proof of Exercise 4.11 a).

Exercise 4.13 (Kurtosis of GARCH(1,1) process)
Let $(X_t)_{t\in\mathbb{Z}}$ be a covariance-stationary GARCH$(1,1)$ process with finite kurtosis κ_Z of the innovation distribution.

a) By using the solution of the GARCH$(1,1)$ defining equations (see MFE (2015, equation (4.27)) show that $(X_t)_{t\in\mathbb{Z}}$ has a finite fourth moment if and only if

$$(\alpha_1+\beta_1)^2 < 1-(\kappa_Z-1)\alpha_1^2,$$

where α_1 and β_1 are the ARCH and GARCH parameters, respectively.

b) Suppose you know that the kurtosis κ_X of the GARCH$(1,1)$ process is equal to the kurtosis of the innovation distribution. Determine κ_X in this case and comment on whether this situation is likely to arise when modelling financial return data.

Exercise 4.14 (GARCH(2,1) and GARCH(1,2) processes)
Let $(X_t)_{t\in\mathbb{Z}}$ be a covariance-stationary GARCH$(2,1)$ process.

a) Derive a recursive formula for $P_tX_{t+h}^2 = \mathbb{E}(X_{t+h}^2 \mid \mathcal{F}_t)$ for $h\geq 1$.

b) Compute the limit $\lim_{h\to\infty} P_tX_{t+h}^2$ and infer the condition for covariance stationarity from this limit.

c) Repeat a) and b) for the GARCH$(1,2)$ process.

Exercise 4.15 (Skewed innovation distributions)
This question concerns the method that is used in the R package `rugarch` to introduce skewness into the innovation distribution of a GARCH process. Let f_X denote a unimodal density that is symmetric around 0 and consider the density f_Y given by

$$f_Y(x) = \frac{2}{\delta+\delta^{-1}}(f_X(\delta x)H(-x) + f_X(\delta^{-1}x)H(x)), \quad x\in\mathbb{R},$$

where $\delta > 0$ is a skewness parameter and $H(x)$ is the Heaviside function defined by $H(x)=0$ for $x<0$, $H(0)=0.5$ and $H(x)=1$ for $x>0$.

a) Explain the intuition behind this construction.

b) Verify that f_Y is a probability density and calculate the distribution function F_Y of f_Y.

c) Let X be a random variable with density f_X and let Y be a random variable with density f_Y. Calculate the kth moment of Y in terms of the kth moment of X, assuming that the latter exists.

d) The distribution of Y must be standardized to have mean zero and unit variance for use as the innovation distribution of a GARCH process. Calculate $\mathbb{E}(Y)$ and $\mathrm{var}(Y)$ in the case where $X\sim \mathrm{N}(0,1)$.

e) Repeat the calculations in the case where $X\sim t_\nu$, $\nu>1$.

Exercise 4.16 (Fitting GARCH models to equity index return data)
Take the set of closing prices of the S&P 500 index from 2006 to 2009 including the financial crisis; see the dataset `SP500` of the R package `qrmdata`.

a) Fit GARCH$(1,1)$ models with standard normal, standardized t and standardized skewed t innovations to the negative log-returns; see the functions `ugarchspec()` and `ugarchfit()` in the R package `rugarch`. Plot the resulting standardized residuals and Q-Q plots against the fitted innovation distributions.
b) Add AR(1), MA(1) and ARMA$(1,1)$ models for the conditional mean and investigate whether this improves the resulting fit of the models.
c) Use your favourite model to calculate an in-sample estimate of volatility and an in-sample estimate of the 99% value-at-risk for the period.
d) Compare the GARCH volatility estimate with an exponentially-weighted moving-average (EWMA) estimate of volatility. Note that this can be carried out in `rugarch` by choosing an IGARCH model in `ugarchspec()` and setting the parameters to have fixed values.

Exercise 4.17 (Fitting GARCH models to weekly equity index return data)
Repeat the analysis of Exercise 4.16 for the weekly returns of the S&P 500 index from 2000 to 2015.

Exercise 4.18 (Fitting GARCH models to foreign-exchange return data)
Repeat the analysis of Exercise 4.16 for daily log-returns of the EUR/USD exchange rate from 2012 to 2015; see the dataset `EUR_USD` of the R package `qrmdata`.

Advanced

Exercise 4.19 (AR(2) process)
Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary and causal solution of the AR(2) equations

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = \varepsilon_t, \quad t \in \mathbb{Z},$$

where $\phi_1 \in \mathbb{R}$, $\phi_2 \neq 0$ and $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0, \sigma_\varepsilon^2)$.

a) Using the general ARMA theory in MFE (2015, p. 103), show that the AR parameters must satisfy the following conditions:
   i) $|\phi_2| < 1$;
   ii) $\phi_2 + \phi_1 < 1$; and
   iii) $\phi_2 - \phi_1 < 1$.

   Note. One can also show that i)–iii) are sufficient.
b) What shape is the set of permissible parameters (ϕ_1, ϕ_2) satisfying i)–iii)?
c) Derive a recurrence relation for the autocorrelation function of $(X_t)_{t\in\mathbb{Z}}$ of the form $\rho(h) = a_1\rho(h-1) + a_2\rho(h-2)$ for some a_1, a_2. What is $\rho(0)$, $\rho(1)$?

Exercise 4.20 (EGARCH(1,1) model)
The exponential GARCH (EGARCH) model is a widely-used alternative to the standard GARCH model. The EGARCH(1, 1) takes the form $X_t = \sigma_t Z_t$ where $(Z_t)_{t\in\mathbb{Z}} \sim \mathrm{SWN}(0,1)$ and the logarithm of the volatility σ_t is updated according to the equations

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1 Z_{t-1} + \gamma_1(|Z_{t-1}| - \mathbb{E}(|Z_{t-1}|)) + \beta_1 \log(\sigma_{t-1}^2), \quad t \in \mathbb{Z},$$

where $\alpha_0, \alpha_1, \beta_1, \gamma_1 \in \mathbb{R}$.

a) Give an economic interpretation for the form of the squared volatility equations.
b) Let $\varepsilon_t = \alpha_1 Z_{t-1} + \gamma_1(|Z_{t-1}| - \mathbb{E}(|Z_{t-1}|))$, $t \in \mathbb{Z}$. Is the process $(\varepsilon_t)_{t\in\mathbb{Z}}$ a strict white noise? Calculate the mean and variance of ε_t, $t \in \mathbb{Z}$, under the assumption that Z_t is symmetrically distributed around 0.
c) For $t \in \mathbb{Z}$, let $Y_t = \log(\sigma_t^2)$ and let ε_t be defined as in b). What kind of process does $(Y_t)_{t\in\mathbb{Z}}$ follow?
d) Give a condition for $(Y_t)_{t\in\mathbb{Z}}$ to be covariance stationary and derive the mean and variance of Y_t, $t \in \mathbb{Z}$.

Exercise 4.21 (GJR-GARCH(1,1) model)
The GJR-GARCH model of Glosten et al. (1993) is another extension of the standard GARCH model. The GJR-GARCH(1, 1) model takes the form $X_t = \sigma_t Z_t$ where $(Z_t)_{t\in\mathbb{Z}} \sim \mathrm{SWN}(0,1)$ and the volatility σ_t is updated according to the equations

$$\sigma_t^2 = \alpha_0 + (\alpha_1 + \gamma_1 I_{\{X_{t-1}<0\}})X_{t-1}^2 + \beta_1\sigma_{t-1}^2, \quad t \in \mathbb{Z},$$

where $\alpha_0 > 0, \alpha_1 + \gamma_1 \geq 0$ and $\beta_1 \geq 0$.

a) Give an economic interpretation for the form of the squared volatility equations.
b) By considering the variance of the process $\mathrm{var}(X_t)$, $t \in \mathbb{Z}$, show that a necessary condition for $(X_t)_{t\in\mathbb{Z}}$ to be covariance stationary is that

$$\alpha_1 + \beta_1 + \gamma_1\theta_Z < 1,$$

where $\theta_Z = \mathbb{E}(I_{\{Z_t<0\}}Z_t^2)$.
c) Show that $\theta_Z = 1/2$ when Z_t has a distribution that is symmetric about 0.
d) Derive an equation for predicting the squared volatility σ_{t+h}^2, $h \geq 1$, given the values of X_t and σ_t^2.

5 Extreme Value Theory

Review

Exercise 5.1 (Maximum domain of attraction)

a) What is the name of the theorem that describes the limiting behaviour of sample maxima?

b) Explain what it means to say that "a distribution function F belongs to the maximum domain of attraction of H_ξ for some $\xi \in \mathbb{R}$".

c) Why is the knowledge that $F \in \mathrm{MDA}(H_\xi)$ for some $\xi \in \mathbb{R}$ so useful for modelling the distribution of maximal losses from F?

Exercise 5.2 (Characterization of the MDA of a GEV for positive shape)

Characterize the maximum domain of attraction of H_ξ for $\xi > 0$. In other words, state precisely which distribution functions F belong to $\mathrm{MDA}(H_\xi)$, $\xi > 0$?

Exercise 5.3 (Block maxima method)

a) Explain the block maxima method for estimating the parameters of a generalized extreme value distribution.

b) Explain why the block maxima method involves a bias-variance trade-off.

Exercise 5.4 (Excess distribution and mean excess function)

a) For a random variable $X \sim F$, how are the excess distribution function over a threshold u and the mean excess function defined?

b) What is the excess distribution function of the (memoryless) exponential distribution $X \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$?

c) What is the mean excess function of the exponential distribution?

Exercise 5.5 (VaR, ES and the mean excess function)

Let $X \sim F$ be continuously distributed with $\mathbb{E}(|X|) < \infty$. Give a formula relating the expected shortfall (ES) of X to the value-at-risk (VaR) of X via the mean excess function.

Exercise 5.6 (Peaks-over-threshold method)

a) What is the name of the theorem that underlies the peaks-over-threshold (POT) method?

b) Explain the POT method for estimating the distribution of excess losses above a threshold.

c) Explain why the POT method involves a bias-variance trade-off.

d) Does the POT method represent a more or a less efficient use of the data on extreme events than the block maxima method?

e) Explain the key ideas in the estimation of tail probabilities based on the POT method.

Exercise 5.7 (Threshold choice in the peaks-over-threshold method)
Figure E.5.1 shows sample mean excess plots for two sets of financial losses.

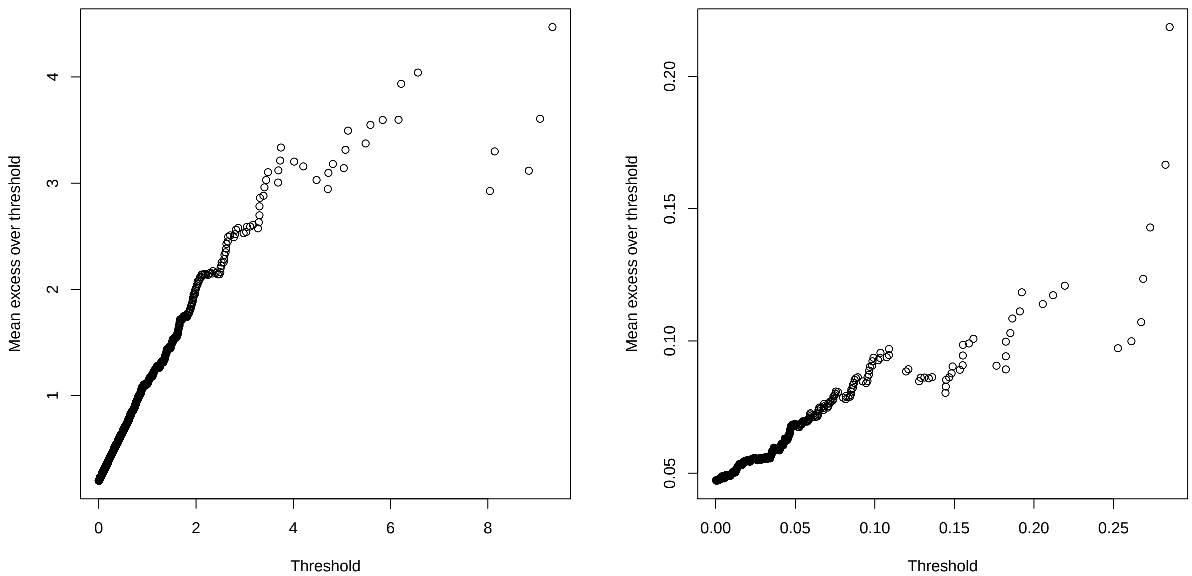

Figure E.5.1 Sample mean excess plots for two sets of financial losses.

a) For each plot in Figure E.5.1 suggest a threshold u which could be used when applying the POT method. Justify your choice.
b) Can you use the plots to give rough estimates of the shape parameter ξ of an approximate GPD model for the excesses over u for the two datasets?

Exercise 5.8 (The tail index and its estimation)

a) Define the tail index of a distribution $F \in \mathrm{MDA}(H_\xi)$ for $\xi > 0$.
b) Describe two possible approaches to estimating the tail index for a suitable dataset.

Exercise 5.9 (The extremal index)
Explain what is meant by the extremal index of a strictly stationary process $(X_i)_{i\in\mathbb{Z}}$.

Basic

Exercise 5.10 (Max-stability of the Fréchet distribution)
A random variable X is *max-stable* if for any $n \in \mathbb{N}$ the maximum $M_n = \max\{X_1, \dots, X_n\}$ of independent copies $X_1, \dots, X_n$ of X satisfies

$$(M_n - d_n)/c_n \stackrel{\mathrm{d}}{=} X$$

for real constants $c_n > 0$ and d_n. By determining appropriate sequences (c_n) and (d_n), prove that the Fréchet distribution with distribution function $F(x) = \exp(-x^{-\theta})$, $x > 0$, $\theta > 0$, is max-stable.

Exercise 5.11 (A continuous distribution for which maxima do not converge to a GEV)
Let X be log-Pareto distributed with distribution function F_X, that is $X = \exp(Y)$ for $Y \sim \mathrm{Pa}(\alpha, \kappa)$; see MFE (2015, Appendix A.2.8). Show that $F_X \notin \mathrm{MDA}(H_\xi)$ for any $\xi \in \mathbb{R}$.

Exercise 5.12 (Double logarithmic tail plot)

a) Suppose you have data $x_1, \dots, x_n$ which are believed to represent realizations of iid random variables with distribution function $F \in \mathrm{MDA}(H_\xi)$ for $\xi > 0$. A popular diagnostic plot for such data is $(x_i, \bar{F}(x_i))$, $i \in \{1, \dots, n\}$, using logarithmic x and y-axes. Explain why.
b) Explain how the plot can be used to construct an estimator of the tail index.

Exercise 5.13 (Uniqueness of limit up to location and scale)
In MFE (2015, Example 5.6) it is shown that if we take the Pareto distribution with distribution function $F(x) = 1 - (\kappa/(\kappa + x))^\alpha$, $x \geq 0$, $\alpha, \kappa > 0$ and the normalizing sequences $c_n = \kappa n^{1/\alpha}/\alpha$ and $d_n = \kappa(n^{1/\alpha} - 1)$ for $n \in \mathbb{N}$, then

$$\lim_{n \to \infty} F^n(c_n x + d_n) = H_{1/\alpha}(x) \; .$$

Suppose we consider the alternative normalizing sequences $c_n = F^{\leftarrow}(1 - 1/n)$ and $d_n = 0$ for $n \in \mathbb{N}$. Show that $F^n(c_n x + d_n)$ converges to $H_{1/\alpha,\mu,\sigma}(x)$ as $n \to \infty$ for some $\mu \in \mathbb{R}$ and $\sigma > 0$ which you should determine.

Exercise 5.14 (GPD with negative shape parameter)
Let X follow a generalized Pareto distribution with shape $\xi < 0$ and scale $\beta > 0$. Derive the distribution of $Y = -\xi X/\beta$.

Exercise 5.15 (Slow variation, regular variation)

a) Show that the following functions are slowly varying at ∞.
   i) $L(x) = 2 + \cos(1/x)$, $x > 0$.
   ii) $L(x) = \log x$, $x > 0$.
   iii) $L(x) = 1/\log x$, $x > 1$.
   iv) $L(x) = \tilde{L}(1/x)$ where $\tilde{L} : [0, \infty) \to [0, \infty)$ is slowly varying at 0.
b) Show that the following functions are regularly varying at ∞ and determine the corresponding index.
   i) $h(x) = x^{-\theta}$, $x > 0$, $\theta > 0$.
   ii) $h(x) = (1 + x)^{-\theta}$, $x > -1$, $\theta > 0$.
   iii) $h(x) = \tilde{h}(1/x)$, $x > 0$, where $\tilde{h}$ is regularly varying at 0 with index $\alpha \in \mathbb{R}$.

Exercise 5.16 (GEV limit for maxima from the GPD)
Show that $G_{\xi,\beta} \in \mathrm{MDA}(H_\xi)$, $\xi \in \mathbb{R}$.

Exercise 5.17 (GEV limit for maxima from the Burr distribution)
Show that the *Burr distribution* with distribution function

$$F(x) = 1 - (\kappa/(\kappa + x^{\theta}))^{\lambda}, \quad x \geq 0,\ \kappa, \theta, \lambda > 0,$$

is in the maximum domain of attraction of H_ξ for some ξ which you should determine in terms of the parameters of the distribution.

Exercise 5.18 (The Weibull distribution as a special case of the GEV)
For $\xi < 0$ the generalized extreme value (GEV) distribution takes the form

$$H_\xi(x) = \exp(-(1 + \xi x)^{-1/\xi}) \quad \text{for all} \quad x < -1/\xi.$$

This case is referred to as a Weibull distribution although it differs from the standard Weibull distribution used in statistics and actuarial science. Let $X \sim H_\xi$ and $Y = 1 + \xi X$.

a) Derive the distribution function of Y and its domain.
b) Verify that Y has a standard Weibull distribution with density

$$f_Y(y) = c\gamma y^{\gamma-1} \exp(-cy^{\gamma}), \quad y > 0,$$

for parameters $c > 0$ and $\gamma > 0$ which you should determine in terms of ξ.

Exercise 5.19 (An alternative formula for the mean excess function)

a) Let $X \sim F$ with $\mathbb{E}(|X|) < \infty$. Show that the mean excess function can be calculated using the formula

$$e(u) = \int_0^{x_F - u} \bar{F}_u(x)\,\mathrm{d}x = \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x)\,\mathrm{d}x, \quad u \in [\bar{x}_F, x_F),$$

where $\bar{x}_F = \inf\{x \in \mathbb{R} : F(x) > 0\}$ denotes the left endpoint of F.

Hint. First show that $x\bar{F}(x) \to 0$ for $x \to x_F$. Conclude that $\lim_{x \to x_F - u} x\bar{F}_u(x) = 0$. Then use the Stieltjes form of the integration by parts formula

$$\int_a^b h(x)\,\mathrm{d}g(x) = \Big[h(x)g(x)\Big]_a^b - \int_a^b g(x)\,\mathrm{d}h(x);$$

see Apostol (1974, p. 144) or Arendt et al. (2002, p. 52).

b) Use the formula in a) to calculate the mean excess function of a random variable X following a generalized Pareto distribution with shape $\xi < 1$ and scale $\beta > 0$.

Exercise 5.20 (Scaling of VaR, ES under a power tail)
Let $X \sim F$ where $F(x) = 1 - c(\kappa + x)^{-\theta}$, $x \geq u$, for some $c > 0$, $\kappa \in \mathbb{R}$, $\theta > 0$ and $u > 0$.

a) Show that, for all α_1 and α_2 satisfying $F(u) < \alpha_1 \leq \alpha_2 < 1$, it holds that

$$\mathrm{VaR}_{\alpha_2}(X) = \left(\frac{1 - \alpha_1}{1 - \alpha_2}\right)^{1/\theta} (\kappa + \mathrm{VaR}_{\alpha_1}(X)) - \kappa.$$

b) Let $\theta > 1$. Show that, for all α_1 and α_2 satisfying $F(u) < \alpha_1 \leq \alpha_2 < 1$, it holds that

$$\mathrm{ES}_{\alpha_2}(X) = \frac{\theta}{\theta - 1}\left(\frac{1-\alpha_1}{1-\alpha_2}\right)^{1/\theta}(\kappa + \mathrm{VaR}_{\alpha_1}(X)) - \kappa.$$

c) Explain why these results are useful from a practical point of view.

Exercise 5.21 (Block maxima method applied to S&P 500 return data)
Imagine you are a risk manager sitting in your office in the evening of Friday 26 September 2008. It has been a turbulent period in the financial market and the S&P500 index has dropped by 3.33% over the course of the week and by 11.82% over the past quarter.

Use statistical software to carry out a block maxima analysis of daily negative log-returns of the S&P500 index since 1960 suing a block size of one year; see the dataset `SP500` of the R package `qrmdata`. In other words, fit a GEV distribution $H_{\boldsymbol{\theta}}$ to the block maxima where $\boldsymbol{\theta} = (\xi, \mu, \sigma)$.

Answer the following questions:

a) Compute the maximum likelihood estimator of the parameter vector $\boldsymbol{\theta}$. Based on the estimated shape parameter, which moments of the fitted GEV are finite?

b) What is the probability that next year's maximum loss exceeds the maximal loss over the past 20 years?

c) What are (approximately) the 10-year and 50-year 260-block return levels $r_{260,10}$ and $r_{260,50}$, respectively?

   Note. The index 260 stands for the approximate number of trading days per year here.

d) What is the return period $k_{260,0.0922}$ of a loss at least as large as 9.22%, as seen on Monday, 29 September 2008?

Exercise 5.22 (Peaks-over-threshold analysis of Bitcoin data)
The objective of this exercise is to use the POT method to analyse extreme negative log-returns on the dollar price of one Bitcoin from 1 January 2014 to 31 December 2017; see the dataset `crypto` of the R package `qrmdata` and the R package `qrmtools` for functions for implementing the POT method.

Work through the following steps:

a) Load the Bitcoin price data and plot it. Also compute the series of negative log-returns and plot these.

b) For the negative log-return data, use a mean excess plot to find a suitable threshold u. Check your choice with a plot of fitted GPD shape parameters against various thresholds.

c) Compute the maximum likelihood estimator of the GPD parameters based on the excess losses over your chosen threshold u. Use a Q-Q plot to verify whether the fitted GPD fits the excess losses well. Also provide a plot of the empirical distribution function of the excess losses overlaid with the fitted GPD to see whether the two match.

d) Use a POT-based tail estimator to compute $\mathrm{VaR}_{0.99}$ and $\mathrm{ES}_{0.975}$ and compare the two values.

e) Plot the empirical tail probabilities $\bar{F}_n(x)$, where F_n denotes the empirical loss distribution function, at all exceedances x of the threshold u and overlay the points with the tail estimator obtained from the POT method; see MFE (2015, equation (5.21)). Also indicate $\mathrm{VaR}_{0.99}$ and $\mathrm{ES}_{0.975}$ in the plot.

Advanced

Exercise 5.23 (Alternative condition for convergence of normalized maxima)

a) Show that

$$F^n(c_n x + d_n) \underset{n\to\infty}{\to} H(x) \quad \text{if and only if} \quad n\bar{F}(c_n x + d_n) \underset{n\to\infty}{\to} -\log H(x).$$

Hint. You may use that $a_n \to a \in \mathbb{R}$ for $n \to \infty$ if and only if $(1 + a_n/n)^n \to \exp(a)$ for $n \to \infty$.

b) Use part a) to show that the exponential distribution function $F(x) = 1 - \exp(-\lambda x)$, $x \geq 0$, $\lambda > 0$, belongs to the maximum domain of attraction of the Gumbel distribution $H_0(x) = \exp(-e^{-x})$, $x \in \mathbb{R}$, for the normalizing sequences $c_n = 1/\lambda$ and $d_n = \log(n)/\lambda$, $n \in \mathbb{N}$.

Exercise 5.24 (Maximum domain of attraction of a log-gamma distribution)

In MFE (2015, Example 16.1) Karamata's Theorem (MFE (2015, Theorem A.7)) is used to show that the Student t distribution is in the maximum domain of attraction of the Fréchet distribution. Use a similar method to analyse the behaviour of normalized maxima for the log-gamma distribution $\mathrm{LGa}(\alpha, \beta)$ with distribution function F and density

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\log x)^{\alpha-1} x^{-\beta-1}, \quad x > 1,\ \alpha > 0,\ \beta > 0.$$

In other words, determine the ξ for which $F \in \mathrm{MDA}(H_\xi)$.

Exercise 5.25 (Maximum domain of attraction under serial dependence)

Suppose $(X_n)_{n\in\mathbb{N}}$ is a sequence of iid random variables with distribution function $F(x) = \sqrt{1 - x^{-\theta}}$, $x \geq 1$, $\theta > 0$. Furthermore, let

$$\begin{aligned} Y_n &= \max\{X_{2n-1}, X_{2n}\}, \quad n \in \mathbb{N}, \\ Z_n &= \max\{X_n, X_{n+1}\}, \quad n \in \mathbb{N}. \end{aligned}$$

a) Clearly, $(Y_n)_{n\in\mathbb{N}}$ is a sequence of iid random variables. Compute the distribution function of Y_1.

b) Let $M_n^Y = \max\{Y_1, \dots, Y_n\}$, $n \in \mathbb{N}$. Find real sequences (c_n) and (d_n), with $c_n > 0$ for all n, and a non-degenerate distribution function H^Y such that $(M_n^Y - d_n)/c_n$ converges in distribution to H^Y.

c) Compute the distribution function of Z_1.

d) Determine whether $(Z_n)_{n\in\mathbb{N}}$ is also a sequence of independent random variables.

e) Let $M_n^Z = \max\{Z_1, \dots, Z_n\}$, $n \in \mathbb{N}$. Using the same sequences (c_n) and (d_n) as in b) show that $(M_n^Z - d_n)/c_n$ converges in distribution to a limit H^Z.

f) How are the limiting distributions H^Y and H^Z related and what is the connection with the extremal index?

Exercise 5.26 (A link between the Poisson, GPD and GEV distributions)
Suppose that over some fixed period we observe a Poisson-distributed number of losses $N \sim \mathrm{Poi}(\lambda)$ for some $\lambda > 0$. Let the loss severities be described by an iid sequence of random variables $(Y_n)_{n\in\mathbb{N}}$ from a GPD with shape $\xi \in \mathbb{R}$ and scale $\beta > 0$. Assume that the loss severities $(Y_n)_{n\in\mathbb{N}}$ are independent of the loss frequency N. Let

$$M_N = \max\{Y_1, \dots, Y_N\}.$$

Show that $M_N \sim H_{\xi,\mu,\sigma}$ where

$$\mu = \begin{cases} \beta(\lambda^\xi - 1)/\xi, & \xi \neq 0, \\ \beta \log \lambda, & \xi = 0, \end{cases} \quad \text{and} \quad \sigma = \beta\lambda^\xi.$$

Exercise 5.27 (Convergence of the excess distribution for a gamma distribution)
Let $X \sim \mathrm{Ga}(2,1)$ be gamma distributed. In solving this exercise it will be seen that the Pickands–Balkema–de Haan Theorem is particularly easy to verify in this case.

a) Compute the excess distribution function F_u over a threshold $u > 0$.

b) Show that

$$\lim_{u\to\infty} \sup_{0<x<\infty} |F_u(x) - G_{0,1}(x)| = 0.$$

c) Using b), identify the maximum domain of attraction to which $\mathrm{Ga}(2,1)$ belongs.

d) Compute $e(u)$ and explain why this implies the same conclusion as in c).

Exercise 5.28 (Shortfall-to-range-shortfall ratio for a GPD tail)
Let L have distribution function $F \in \mathrm{MDA}(H_\xi)$ and assume that for some threshold $u \in [\bar{x}_F, x_F)$ the excess distribution exactly satisfies $F_u(x) = G_{\xi,\beta}(x)$, $x \in [0, x_F - u)$, $\xi \in \mathbb{R}$, where $\bar{x}_F = \inf\{x \in \mathbb{R} : F(x) > 0\}$ denotes the left endpoint of F.

a) For $0 < \alpha_1 < \alpha_2 < 1$, compute the *range expected shortfall*

$$\mathrm{ES}_{\alpha_1,\alpha_2}(L) = \frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} \mathrm{VaR}_z(L)\,\mathrm{d}z.$$

b) For $\xi < 1$, derive $\mathrm{ES}_\alpha(L)$.

c) For $\xi < 1$, show that for all $\gamma > 1$,

$$\lim_{\alpha \to 1-} \frac{\mathrm{ES}_\alpha(L)}{\mathrm{ES}_{\alpha, 1-(1-\alpha)^\gamma}(L)} = 1$$

and interpret this result from a risk management perspective.

Exercise 5.29 (Probability-weighted moments for the GEV distribution)
If they exist, the *probability-weighted moments* of a random variable X with distribution function F are defined by

$$M_{p,r,s} = \mathbb{E}(X^p F(X)^r (1 - F(X))^s), \quad p, r, s \in \mathbb{R}.$$

For continuous F and $U \sim \mathrm{U}(0,1)$, this implies that $M_{p,r,s} = \mathbb{E}(F^{\leftarrow}(U)^p U^r (1-U)^s)$. Consider the probability-weighted moments given by $\beta_r = M_{1,r,0}$ and assume that $F = H_{\xi,\mu,\sigma}$ for some $\xi < 1$.

a) Compute β_r for $r > -1$.

   Hint. You may use the fact that $\int_0^\infty \log(x)\exp(-x)\,\mathrm{d}x = -\gamma$, where γ denotes the Euler–Mascheroni constant.

b) Hence compute β_0, $2\beta_1 - \beta_0$ and $(3\beta_2 - \beta_0)/(2\beta_1 - \beta_0)$.

c) Based on b) and estimates $\beta_{r,n}$ of β_r, design an algorithm for computing estimates of ξ, μ, σ.

Exercise 5.30 (Probability-weighted moments for the GPD)
In this exercise we apply the method of probability-weighted moments as introduced in Exercise 5.29 to the estimation of a generalized Pareto distribution $F = G_{\xi,\beta}$ under the assumption that $\xi < 1$. For the GPD it proves easier to consider the probability-weighted moments given by $\alpha_s = M_{1,0,s}$.

a) Compute α_s for $s \geq 0$.

b) Hence compute $2 - \alpha_0/(\alpha_0 - 2\alpha_1)$ and $2\alpha_0\alpha_1/(\alpha_0 - 2\alpha_1)$.

c) Based on b) and estimates $\alpha_{s,n}$ of α_s, design an algorithm for computing estimates of ξ, β.

6 Multivariate Models

Review

Exercise 6.1 (Sums of normal random variables)
Is it true to state that if $L_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $L_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ then $L_1 + L_2 \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$? Explain your answer.

Exercise 6.2 (Independence of uncorrelated normal variables)
Let $\boldsymbol{X} \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$. Explain why we can infer that if $X_1, \dots, X_d$ are uncorrelated they are also independent.

Exercise 6.3 (Stochastic representation for normal variance mixtures)
Let $\boldsymbol{X} \sim \mathrm{M}_d(\boldsymbol{\mu}, \Sigma, \hat{H})$ be a random vector with a normal variance mixture distribution. Let W be a random variable with Laplace-Stieltjes transform $\hat{H}$. Assume that Σ is positive definite with Cholesky decomposition $\Sigma = AA'$.

a) Give a stochastic representation for $\boldsymbol{X}$ which shows how $\boldsymbol{X}$ may be generated starting from W and a vector of independent standard normal random variables $\boldsymbol{Z} = (Z_1, \dots, Z_d)'$.
b) Assuming $X_1, \dots, X_d$ have finite variances, use this representation to explain why uncorrelatedness of $X_1, \dots, X_d$ does not imply their independence, unless W is constant almost surely.

Exercise 6.4 (Normal variance mixtures as elliptical distributions)
Using the stochastic representation in Exercise 6.3, show that $\boldsymbol{X} \sim \mathrm{M}_d(\boldsymbol{\mu}, \Sigma, \hat{H})$ has an elliptical distribution.

Exercise 6.5 (Distinguishing distributions by scatterplots of random samples)
Figure E.6.1 shows random samples from nine bivariate distributions. Match the pictures to the following distributions:

a) a normal distribution;
b) a t distribution with $\nu = 3.5$ degrees of freedom and positive correlation;
c) a t distribution with $\nu = 3.5$ degrees of freedom and negative correlation;
d) a t distribution with $\nu = 3.5$ degrees of freedom and zero correlation;
e) an elliptical distribution with discrete radial part;
f) an elliptical distribution with bounded radial part;
g) a mixture of a normal and a singular distribution;

h) the distribution of $(X_1, X_2(1 - 2I_{\{X_1<0\}}))$ where (X_1, X_2) is jointly normal; and
i) a normal mean-variance mixture.

Exercise 6.6 (Advantages of generalized hyperbolic distributions)
Describe some of the properties and features of generalized hyperbolic distributions which make them appealing for applications in financial risk management.

Exercise 6.7 (Different kinds of factor models)
Explain the key differences between macroeconomic, fundamental and statistical factor models.

Basic

Exercise 6.8 (Independence of uncorrelated Bernoulli variables)
Let (X_1, X_2) follow a bivariate Bernoulli distribution with given joint probabilities $\mathbb{P}(X_1 = k,\ X_2 = l) = p_{kl}$ for $k, l \in \{0, 1\}$. Show that X_1 and X_2 are uncorrelated if and only if they are independent.

Exercise 6.9 (Calculating the correlation coefficient from a joint density)
Consider a bivariate distribution function F with density

$$f(x_1, x_2) = \theta(\theta+1)(x_1 + x_2 - 1)^{-(\theta+2)}, \quad x_1, x_2 \in [1, \infty),\ \theta > 0.$$

a) Derive F.
b) Derive the marginal distribution functions F_1, F_2 of F.
c) Compute the correlation coefficient ρ corresponding to F. What assumption must we make about θ in order that ρ exists?

Exercise 6.10 (Cholesky decomposition)
Let Σ be the positive-definite matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

for $\sigma_1 > 0$, $\sigma_2 > 0$ and $\rho \in (-1, 1)$. Show that Σ has a unique Cholesky factor A. In other words, show that there exists a unique lower-triangular matrix A with positive diagonal entries satisfying $\Sigma = AA'$.

Exercise 6.11 (Density of multivariate normal)
Let $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ where $Z_1, \dots, Z_d$ are iid $\mathrm{N}(0,1)$ random variables. Let $\boldsymbol{X} = \boldsymbol{\mu} + A\boldsymbol{Z}$ where $\boldsymbol{\mu} \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d\times d}$ with $\operatorname{rank} A = d$. Clearly $\boldsymbol{X} \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$ where $\Sigma = AA'$.

a) Write down the joint density $f_{\boldsymbol{Z}}(\boldsymbol{z})$ of $\boldsymbol{Z}$.

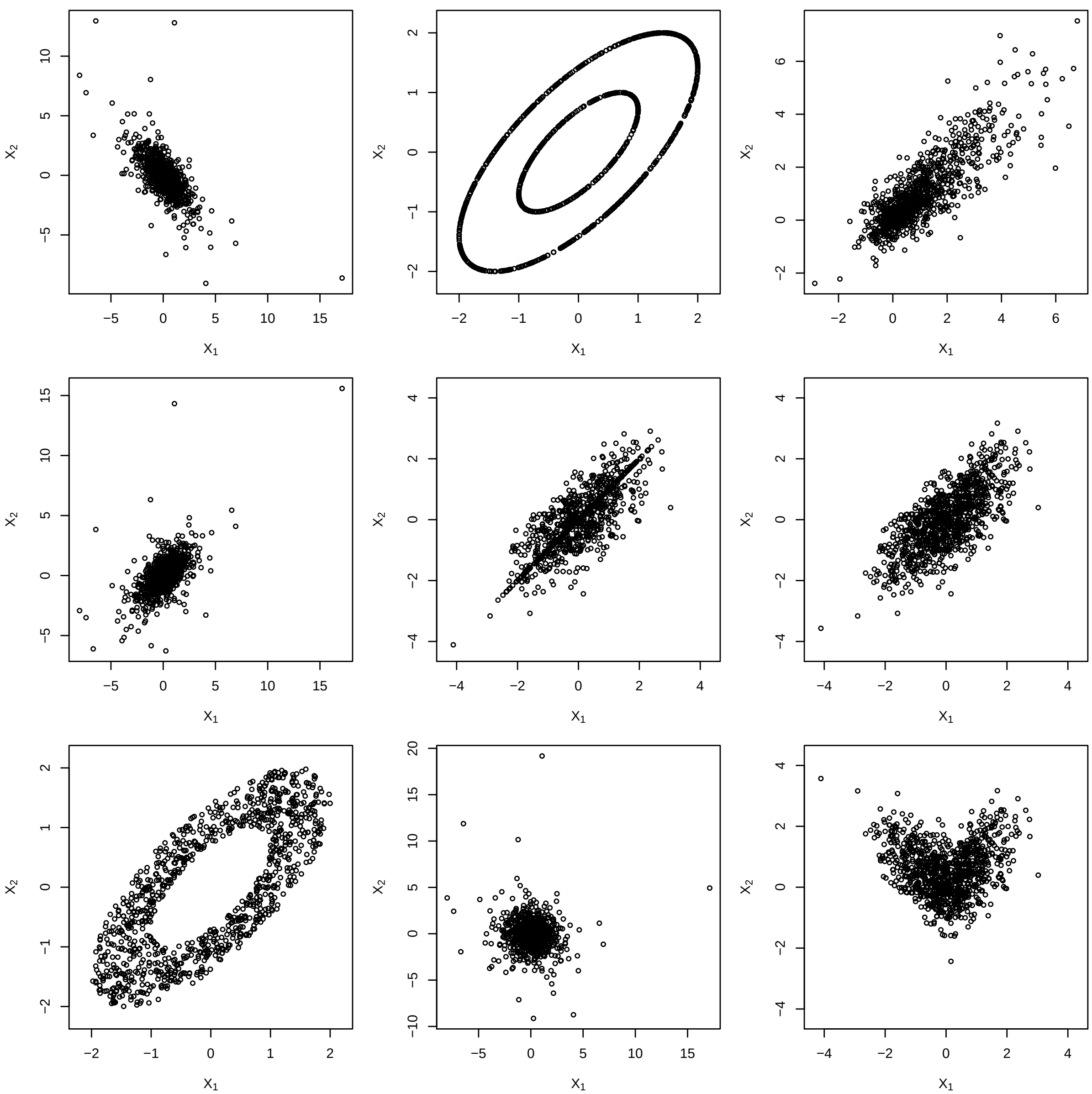

Figure E.6.1 1000 iid samples from various bivariate distributions.

b) Hence, by considering the transformation $T(\boldsymbol{z}) = \boldsymbol{\mu} + A\boldsymbol{z}$ and applying the method of density transformation, show that the density of $\boldsymbol{X}$ is given by

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}\sqrt{\det\Sigma}} \exp\Big(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big), \quad \boldsymbol{x} \in \mathbb{R}^d.$$

Exercise 6.12 (Conditional distribution of bivariate normal)
Let $(X_1, X_2) \sim \mathrm{N}_2(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and assume that $\mathrm{var}(X_i) = \sigma_j^2 > 0$ for $j \in \{1, 2\}$ and $\rho(X_1, X_2) = \rho \in (-1, 1)$. By using the density derived in Exercise 6.11, show that

$$X_2 \,|\, X_1 = x_1 \sim \mathrm{N}\Big(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1),\ \sigma_2^2(1-\rho^2)\Big).$$

Exercise 6.13 (Covariance matrix of a mixture distribution)
Let $\boldsymbol{X} = f(W, \boldsymbol{Z})$ be a random vector whose distribution depends on a random vector $\boldsymbol{Z}$ and an independent, non-negative scalar variable W, such as in the case of a normal variance mixture or a normal mean-variance mixture distribution. Assume the distributions of $\boldsymbol{Z}$ and W are chosen in such a way that the covariance matrix $\mathrm{cov}(\boldsymbol{X})$ exists.

a) Derive the formula $\mathrm{cov}(\boldsymbol{X}) = \mathbb{E}(\mathrm{cov}(\boldsymbol{X} \,|\, W)) + \mathrm{cov}(\mathbb{E}(\boldsymbol{X} \,|\, W))$ in this setup.
b) Calculate $\mathrm{cov}(\boldsymbol{X})$ when $\mathrm{cov}(\boldsymbol{X} \,|\, W) = W\Sigma$ and $\mathbb{E}(\boldsymbol{X} \,|\, W) = \boldsymbol{\mu} + W\boldsymbol{\gamma}$ for a covariance matrix Σ and parameter vectors $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$.

Exercise 6.14 (Normal variance mixture with discrete mixing variable)
Let W be discrete with $\mathbb{P}(W = 1) = p$ and $\mathbb{P}(W = 4) = 1 - p$ for some $p \in (0, 1)$. Furthermore, let Z_1 and Z_2 be iid $\mathrm{N}(0, 1)$ variables, which are also independent of W. For $a \in \mathbb{R}$, let

$$X_1 = \sqrt{W}(Z_1 + 2Z_2) \quad \text{and} \quad X_2 = a + \sqrt{W}Z_2.$$

a) Show that $\boldsymbol{X} = (X_1, X_2)$ has a normal variance mixture distribution.
b) Compute the covariance matrix of $\boldsymbol{X}$ for $p = 2/3$.
c) For $p \in (0, 1)$, find the distribution function of $X_1 + X_2$.

Exercise 6.15 (Normal variance mixture with Pareto Type I mixing variable)
Let $W \sim \mathrm{Pa}(\theta)$, $\theta > 1$, with distribution function $F(x) = 1 - x^{-\theta}$, $x \geq 1$, and let Z_1 and Z_2 be iid $\mathrm{N}(0, 1)$ variables, which are also independent of W. Furthermore, let

$$X_1 = \sqrt{W}(Z_1 + Z_2) \quad \text{and} \quad X_2 = \sqrt{W}(Z_1 - Z_2).$$

a) Show that $\boldsymbol{X} = (X_1, X_2)$ has a normal variance mixture distribution.
b) Compute the covariance matrix of $\boldsymbol{X}$ and show that X_1, X_2 are uncorrelated.

Exercise 6.16 (Diversification and the single-factor model)
Suppose that financial losses $X_1, \ldots, X_d$ on d different assets are described by the single-factor model

$$X_j = \rho Z_0 + \sqrt{1-\rho^2}Z_j, \quad j \in \{1, \ldots, d\},$$

where $\rho \in (0, 1)$ and $Z_0, Z_1, \ldots, Z_d$ are iid $\mathrm{N}(0, \sigma^2)$ random variables.

a) Calculate the covariance matrix of $\boldsymbol{X} = (X_1, \dots, X_d)$.
b) Determine the joint distribution of $\boldsymbol{X}$.
c) Calculate $\mathrm{VaR}_\alpha(X_1)$ and $\mathrm{VaR}_\alpha(\bar{X}_d)$ for $\bar{X}_d = \frac{1}{d}\sum_{j=1}^d X_j$ and $\alpha \in (0,1)$.
d) Does investing a capital of $1/d$ in each of $X_1, \dots, X_d$ decrease VaR_α compared to investing the whole capital 1 in X_1? Can this diversification effect be significantly improved by increasing d if d is already large? Explain your answers.

Exercise 6.17 (VaR of the sum of jointly normal variables)

a) Let (X_1, X_2) be normally distributed with correlation $\rho \in [-1,1]$. Compute $\mathrm{VaR}_\alpha(X_1+X_2)$.
b) Let $\boldsymbol{X} = (X_1, \dots, X_d) \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$. Compute $\mathrm{VaR}_\alpha(X_1 + \dots + X_d)$.

Exercise 6.18 (Contrasting models for uncorrelated normal variables)

Let $(X_1, X_2) \sim \mathrm{N}_2(\boldsymbol{0}, I_2)$ and $(Y_1, Y_2) = (X_1, VX_1)$ where V is independent of X_1 with $\mathbb{P}(V=-1) = \mathbb{P}(V=1) = 1/2$.

a) Show that (X_1, X_2) and (Y_1, Y_2) have the same margins.
b) Show that (X_1, X_2) and (Y_1, Y_2) are both uncorrelated pairs.
c) Compute the distribution functions of $X_1 + X_2$ and $Y_1 + Y_2$.
d) Compute $\mathrm{VaR}_\alpha(X_1+X_2)$ and $\mathrm{VaR}_\alpha(Y_1+Y_2)$ for $\alpha = 0.9$ and $\alpha = 0.99$ and interpret the results.

Exercise 6.19 (VaR-minimizing portfolio for a normal variance mixture)

Consider the normal variance mixture $\boldsymbol{X} = \sqrt{W}\boldsymbol{Y}$ for $\boldsymbol{Y} \sim \mathrm{N}_d(\boldsymbol{0}, \Sigma)$, where $W \geq 0$ is a random variable independent of $\boldsymbol{Y}$ and Σ is a positive definite covariance matrix.

a) Determine the distribution of $\boldsymbol{a}'\boldsymbol{Y}$ for $\boldsymbol{a} \in \mathbb{R}^d\backslash\{\boldsymbol{0}\}$ and give a stochastic representation for $\boldsymbol{a}'\boldsymbol{Y}$ in terms of a standard normal variable $Z \sim \mathrm{N}(0,1)$; recall that this stochastic representation is a representation of the form $\boldsymbol{a}'\boldsymbol{Y} \overset{\mathrm{d}}{=} f(Z)$, for some function f, which shows how the distribution of $\boldsymbol{a}'\boldsymbol{Y}$ may be stochastically simulated using Z.
b) Using a) show that, for any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d\backslash\{\boldsymbol{0}\}$, we can write $\boldsymbol{a}'\boldsymbol{X} \overset{\mathrm{d}}{=} c_{\boldsymbol{a},\boldsymbol{b}}\boldsymbol{b}'\boldsymbol{X}$ and derive a formula for the constant $c_{\boldsymbol{a},\boldsymbol{b}} \in \mathbb{R}$.
c) Using b) derive a formula for $\mathrm{VaR}_\alpha(\boldsymbol{a}'\boldsymbol{X})$ in terms of $\mathrm{VaR}_\alpha(X_1)$.
d) Suppose $\Sigma = \left(\begin{smallmatrix} 2 & 1/2 \\ 1/2 & 1 \end{smallmatrix}\right)$ and $\boldsymbol{a} = (a_1, 1-a_1)$. For $\alpha \in (1/2, 1)$, determine the value of a_1 that minimizes $\mathrm{VaR}_\alpha(\boldsymbol{a}'\boldsymbol{X})$ and give the corresponding value of $\mathrm{VaR}_\alpha(\boldsymbol{a}'\boldsymbol{X})$.

Exercise 6.20 (Linear combinations of t random variables)

a) Let $\boldsymbol{X} = (X_1, \dots, X_d) \sim t_d(\nu, \boldsymbol{\mu}, \Sigma)$. Use characteristic functions to show that the distribution of the linear combination $\lambda_0 + \boldsymbol{\lambda}'\boldsymbol{X} = \lambda_0 + \sum_{j=1}^d \lambda_j X_j$ is univariate t and identify the corresponding parameters.
b) Now let $X_1, \dots, X_d$ be iid random variables with a standard univariate t distribution with ν degrees of freedom, that is $t_1(\nu, 0, 1)$ in notation of part a). Calculate the characteristic function of $\lambda_0 + \sum_{j=1}^d \lambda_j X_j$ in this case and explain why this is not a special case of a).

Exercise 6.21 (Linear combinations of elliptical random vectors)
Let $\boldsymbol{X} \sim \mathrm{E}_d(\boldsymbol{\mu}, \Sigma, \psi)$, $B \in \mathbb{R}^{k\times d}$ and $\boldsymbol{b} \in \mathbb{R}^k$. Use characteristic functions to show that $\boldsymbol{Y} = B\boldsymbol{X} + b \sim \mathrm{E}_k(B\boldsymbol{\mu} + \boldsymbol{b}, B\Sigma B', \psi)$, that is show that linear combinations of elliptically distributed random vectors are elliptically distributed with the same type of distribution determined by ψ.

Exercise 6.22 (Convolutions of elliptical distributions)
Let $\boldsymbol{X} \sim \mathrm{E}_d(\boldsymbol{\mu}, \Sigma, \psi)$ and $\boldsymbol{Y} \sim \mathrm{E}_d(\tilde{\boldsymbol{\mu}}, c\Sigma, \tilde{\psi})$ for some $c > 0$ be independent. Use characteristic functions to show that

$$\boldsymbol{X} + \boldsymbol{Y} \sim \mathrm{E}_d(\boldsymbol{\mu} + \tilde{\boldsymbol{\mu}}, \Sigma, \bar{\psi})$$

and derive the form of the characteristic generator $\bar{\psi}$.

Exercise 6.23 (Fitting generalized hyperbolic distributions to equity return data)
Take the daily log-return data from 2005 to 2012 for the following 10 components of the Dow Jones index: Apple, Cisco, Disney, IBM, Intel, McDonald's, Microsoft, Nike, Proctor & Gamble, and Walmart. Carry out the following analyses:

a) Fit univariate t, NIG, hyperbolic, GH and VG distributions (both symmetric and asymmetric) to the daily log-returns for each stock. In other words, repeat the analysis of MFE (2015, Example 6.14) for these data. Summarize the results.
b) Fit multivariate t, NIG, hyperbolic, GH and VG distributions (both symmetric and asymmetric) to the 10-dimensional data. In other words, repeat the analysis of MFE (2015, Example 6.15) for these data. Summarize the results.
c) Assume you have invested an equal amount of money in each of these stocks. For the multivariate model that you find best, plot the implied distribution of daily portfolio returns.

The R package `ghyp` contains code to estimate all of these distributions. For the data see the dataset `DJ_const` of the R package `qrmdata`.

Exercise 6.24 (Fitting a one-factor model to equity return data)
Take the daily return data used in Exercise 6.23 and the series of daily log-returns for the Dow Jones 30 index for the same period.

a) Fit a single-index model and use it to estimate the betas of the 10 selected Dow Jones stocks. In other words, carry out an analysis using similar steps to MFE (2015, Example 6.34).
b) Comment on the quality of the resulting factor model.

See the dataset `DJ` of the R package `qrmdata` for daily values of the Dow Jones 30 index.

Exercise 6.25 (Principal components analysis of equity return data)
Take the daily return data used in Exercise 6.23.

a) Carry out a principal components analysis and determine the proportion of the total variability of the data that is accounted for by the first two and the first three principal components. In other words, carry out an analysis using similar steps to MFE (2015, Example 6.36).

b) Determine the factor loadings for the first two principal components and comment on whether they have clear interpretations.

c) Construct time series of the first two principal components and plot them.

Advanced

Exercise 6.26 (Equicorrelation matrix)

Let $\boldsymbol{X} = (X_1, \dots, X_d)$ with $\sigma_j^2 = \mathrm{var}(X_j) \in (0, \infty)$, for $j = 1, \dots, d$, and a correlation matrix P which is an equicorrelation matrix with off-diagonal elements equal to ρ.

a) Show that P is positive semi-definite.

b) Show that $\rho \geq -1/(d-1)$.

Exercise 6.27 (On the invertibility of the sample covariance matrix)

Let $S_u \in \mathbb{R}^{d \times d}$ be the unbiased sample covariance matrix computed from a random sample $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ of dimension d.

a) Show that S_u can be written as

$$S_u = \frac{1}{n-1} X'(I_n - J_n/n)X,$$

where $X = (\boldsymbol{X}_1', \dots, \boldsymbol{X}_n')' \in \mathbb{R}^{n \times d}$ and $J_n = \mathbf{1}\mathbf{1}' \in \mathbb{R}^{n \times n}$ for $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$.

b) Show that $\mathrm{rank}(I_n - J_n/n) = n - 1$.

c) Hence show that S_u has rank at most $\min\{n-1, d\}$ and conclude that S_u is not invertible if $n \leq d$.

Hint. You may use that $\mathrm{rank}(AB) \leq \min\{\mathrm{rank}\, A, \mathrm{rank}\, B\}$ and other identities for the rank of matrices.

Exercise 6.28 (Characterization of the multivariate normal distribution)

a) Show that two d-dimensional random vectors $\boldsymbol{X}, \boldsymbol{Y}$ are equal in distribution if and only if $\boldsymbol{a}'\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{a}'\boldsymbol{Y}$ for all $\boldsymbol{a} \in \mathbb{R}^d$.

b) Show that $\boldsymbol{X} \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$ if and only if $\boldsymbol{a}'\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{a}'\boldsymbol{\mu}, \boldsymbol{a}'\Sigma\boldsymbol{a})$ for all $\boldsymbol{a} \in \mathbb{R}^d$.

Exercise 6.29 (Conditional distribution of multivariate normal)

Let $\boldsymbol{X} \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$ with $\mathrm{rank}\, \Sigma = d$. For some $k \in \{1, \dots, d-1\}$, write $\boldsymbol{X} = (\boldsymbol{X}_1', \boldsymbol{X}_2')$, where $\boldsymbol{X}_1 \in \mathbb{R}^k$, $\boldsymbol{X}_2 \in \mathbb{R}^{d-k}$. Furthermore, let $\boldsymbol{\mu} = (\boldsymbol{\mu}_1', \boldsymbol{\mu}_2')$ and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Show that

$$\boldsymbol{X}_2 \,|\, \boldsymbol{X}_1 = \boldsymbol{x}_1 \sim \mathrm{N}_{d-k}(\boldsymbol{\mu}_{2.1}, \Sigma_{22.1}),$$

where $\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1)$ and $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.
Hint. Consider $\boldsymbol{Y} = \boldsymbol{X}_2 + A\boldsymbol{X}_1$ for $A = -\Sigma_{21}\Sigma_{11}^{-1}$.

Exercise 6.30 (Density of multivariate t)

Let $\boldsymbol{X} = \boldsymbol{\mu} + \sqrt{W} A \boldsymbol{Z}$ for $\boldsymbol{\mu} \in \mathbb{R}^d$, $A \in \mathbb{R}^{d \times d}$ with rank $A = d$ and $\boldsymbol{Z} = (Z_1, \dots, Z_d) \sim \mathrm{N}_d(\boldsymbol{0}, I_d)$ independently of $W \sim \mathrm{Ig}(\nu/2, \nu/2)$ for $\nu > 0$. In other words, $\boldsymbol{X} \sim t_d(\nu, \boldsymbol{\mu}, \Sigma)$, where $\Sigma = AA'$.

a) Show that $\boldsymbol{X}$ has density

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\Gamma((\nu + d)/2)}{\Gamma(\nu/2)(\pi\nu)^{d/2}\sqrt{\det \Sigma}} \left(1 + \frac{(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})}{\nu}\right)^{-(\nu+d)/2}, \quad \boldsymbol{x} \in \mathbb{R}^d.$$

b) By using Stirling's Formula, that is $\Gamma(x + 1) \sim \sqrt{2\pi x}(x/e)^x$ for $x \to \infty$, show that $f_{\boldsymbol{X}}$ converges to the density of $\mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$ as $\nu \to \infty$.

Exercise 6.31 (Conditional distribution of bivariate t)

Let $(X_1, X_2) \sim t_2(\nu, \boldsymbol{\mu}, \Sigma)$ with $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ for $\sigma_1, \sigma_2 > 0$ and $\rho \in (-1, 1)$. By using the density of a multivariate t distribution, see Exercise 6.30, show that

$$X_2 \,|\, X_1 = x_1 \sim t\left(\nu + 1,\ \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1),\ \sigma_2^2(1 - \rho^2)\frac{\nu + ((x_1 - \mu_1)/\sigma_1)^2}{\nu + 1}\right).$$

Exercise 6.32 (An interpretation of the principal components transform)

Let $\boldsymbol{X}$ be a random vector with $\mathbb{E}(\boldsymbol{X}) = \boldsymbol{0}$ and $\operatorname{cov}(\boldsymbol{X}) = \Sigma$. Let $\boldsymbol{Y}$ be the vector given by the principal components transform of $\boldsymbol{X}$. In other words, $\boldsymbol{Y} = \Gamma'\boldsymbol{X}$ where $\Sigma = \Gamma\Lambda\Gamma'$, the columns of Γ contain the orthonormal eigenvectors of Σ and $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_d)$ where $\lambda_1 \geq \dots \geq \lambda_d \geq 0$ denote the sorted eigenvalues of Σ. The jth principal component Y_j is thus $Y_j = \boldsymbol{\gamma}_j'\boldsymbol{X}$, where $\boldsymbol{\gamma}_j$ is the jth column of Γ (the eigenvector corresponding to the jth largest eigenvalue of Σ), also known as jth vector of loadings or jth principal axis.

a) Show that the first principal component Y_1 of $\boldsymbol{X}$ satisfies

$$\operatorname{var}(Y_1) = \max_{\boldsymbol{a} \in \mathbb{R}^d,\ \|\boldsymbol{a}\| = 1} \operatorname{var}(\boldsymbol{a}'\boldsymbol{X}),$$

that is Y_1 is the standardized linear combination $\boldsymbol{a}'\boldsymbol{X}$ of $\boldsymbol{X}$ with maximal variance among all linear combinations.

b) Show that the second principal component Y_2 of $\boldsymbol{X}$ satisfies

$$\operatorname{var}(Y_2) = \max_{\substack{\boldsymbol{a} \in \mathbb{R}^d,\ \|\boldsymbol{a}\| = 1, \\ \boldsymbol{a}'\boldsymbol{\gamma}_1 = 0}} \operatorname{var}(\boldsymbol{a}'\boldsymbol{X}),$$

that is Y_2 is the standardized linear combination $\boldsymbol{a}'\boldsymbol{X}$ of $\boldsymbol{X}$ with maximal variance among all linear combinations orthogonal to the first principal axis.

Hint. Apply the method of Lagrange multipliers. It may also be useful to use the facts that $\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{a}'\boldsymbol{x} = \frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}'\boldsymbol{a} = \boldsymbol{a}$ and $\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}'A\boldsymbol{x} = (A + A')\boldsymbol{x}$.

7 Copulas and Dependence

Review

Exercise 7.1 (The concept of a copula)
Give a definition of a copula.

Exercise 7.2 (The importance of Sklar's Theorem)
State Sklar's Theorem, giving both the main statement and its converse. Explain the relevance of the result from the point of view of studying dependencies between random variables and constructing multivariate models for applications.

Exercise 7.3 (On the importance of the three defining properties of copulas)
Show that the function $W(\boldsymbol{u}) = \max\{\sum_{j=1}^{d} u_j - d + 1, 0\}$ fulfills the following properties but is not a copula for $d \geq 3$:

i) $W(u_1, \dots, u_d) = 0$ if $u_j = 0$ for any j;

ii) $W(1, \dots, 1, u_j, 1, \dots, 1) = u_j$ for all $j \in \{1, \dots, d\}$ and $u_j \in [0, 1]$; and

iii) W is increasing in each component.

Exercise 7.4 (Special cases of normal variance mixture copulas)
Consider a d-dimensional normal variance mixture copula (such as Gauss or t) with off-diagonal entries of the correlation matrix P all equal to $\rho \in [-1, 1]$. What models do we obtain in the following special cases?

a) $\rho = 1$

b) $\rho = 0$

c) $\rho = -1$ and $d = 2$

d) $\rho = -1$ and $d \geq 3$

Exercise 7.5 (Correlation and rank correlation)
Describe some drawbacks of Pearson's linear correlation coefficient that are not shared by rank correlation coefficients (such as Spearman's rho and Kendall's tau).

Exercise 7.6 (Maximal VaR of the sum)
Consider the class of joint distributions of the pair (X_1, X_2) when the marginal distributions are fixed and the copula is allowed to vary. Assume that the marginal distributions are such that X_1 and X_2 have finite variances. For each of the following statements, state with reasons whether they are true or false.

a) $\rho(X_1, X_2)$ is maximal when (X_1, X_2) has the comonotonicity copula M.
b) The variance of $X_1 + X_2$ is maximal when $\rho(X_1, X_2)$ is maximal.
c) For fixed $\alpha \in (0,1)$, $\mathrm{VaR}_\alpha(X_1 + X_2)$ is maximal when $\rho(X_1, X_2)$ is maximal.

Exercise 7.7 (VaR for the sum of two normal risks)
A risk manager makes the following assertion: "If X_1 and X_2 are normally distributed risks with correlation ρ, then I know the value of $\mathrm{VaR}_\alpha(X_1 + X_2)$ for any α". Discuss whether you agree.

Exercise 7.8 (Distinguishing meta-distributions from scatterplots)
All the samples shown in Figure E.7.1 come from multivariate distributions with standard normal margins and differing underlying copulas. Match the pictures to the following copulas.

a) a Gauss copula;
b) a t copula with $\nu = 3.5$ degrees of freedom and positive correlation parameter;
c) a t copula with $\nu = 3.5$ degrees of freedom and negative correlation parameter;
d) a t copula with $\nu = 3.5$ degrees of freedom and zero correlation parameter;
e) a Gumbel copula;
f) a Clayton copula;
g) an elliptical copula with bounded radial part;
h) a copula with singular component;
i) the copula of a normal mean-variance mixture.

Exercise 7.9 (Estimation of a copula)
Explain how a copula can be estimated from multivariate data without making specific parametric assumptions about the marginal distributions of the data.

Basic

Exercise 7.10 (Extracting the copula from a joint distribution)
Suppose that d financial losses have the joint distribution function

$$F(\boldsymbol{x}) = \exp\biggl(-\biggl(\sum_{j=1}^{d} x_j^{-\theta}\biggr)^{1/\theta}\biggr), \quad \boldsymbol{x} \in (0,\infty)^d,\ \theta \in [1,\infty).$$

a) Derive the marginal distributions of F and give their common name.
b) Hence derive the copula C of F and give its common name.
c) Assume that $d = 2$ and $\theta = 2$. Calculate the probability that both losses jointly take values above their 95th percentiles and show that this is almost exactly 12 times larger than the case where the losses are independent.

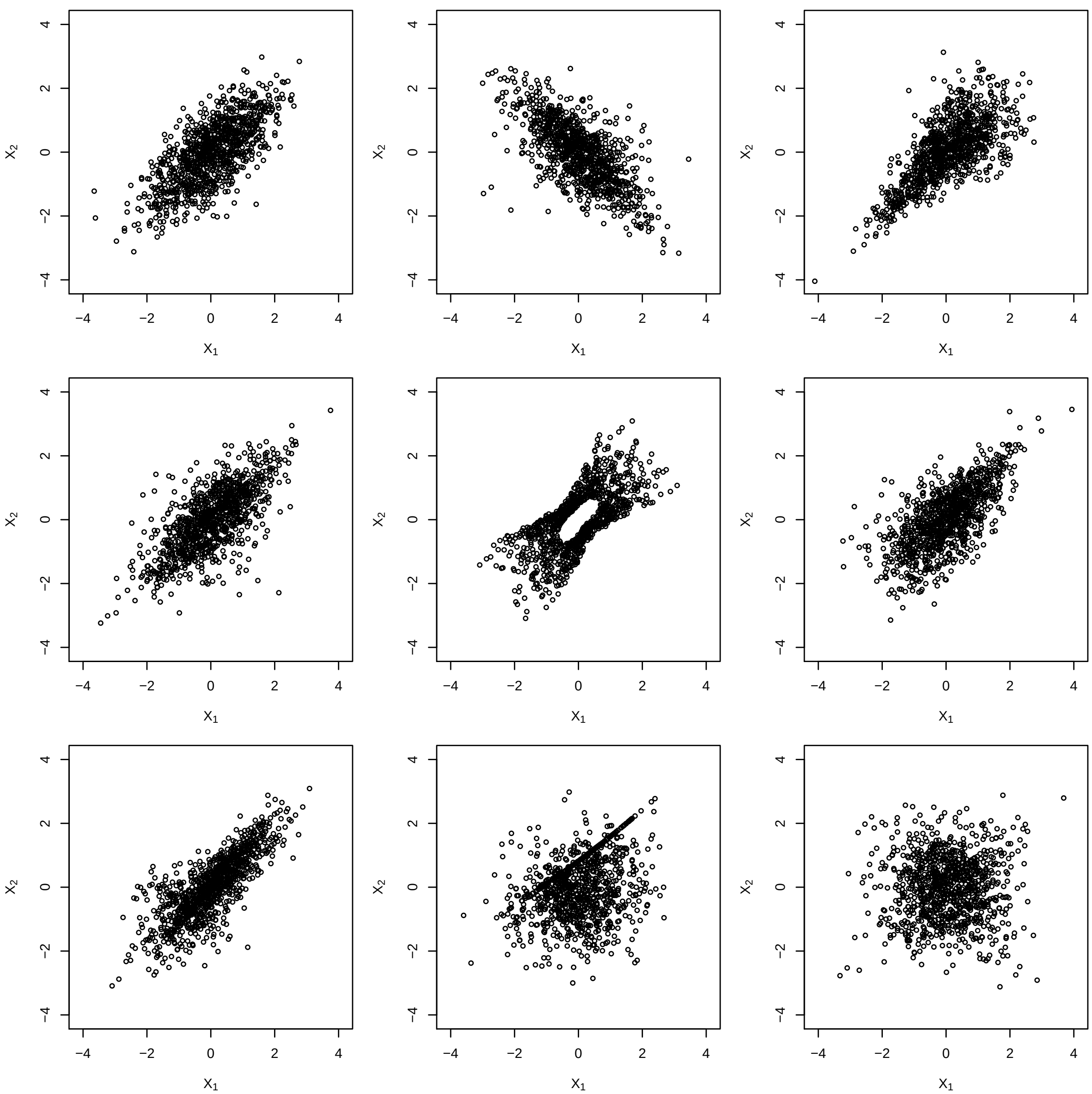

Figure E.7.1 1000 iid samples from various bivariate distributions with $\mathrm{N}(0, 1)$ margins.

Exercise 7.11 (Copula of a normal random variable and its absolute value)
Derive the copula of $Z \sim \mathrm{N}(0,1)$ and $|Z|$.

Exercise 7.12 (Minimal correlation for standard exponential random variables)
Let (X_1, X_2) be a pair of standard exponential random variables with the countermonotonicity copula and let U be a standard uniform random variable.

a) Find a stochastic representation of the form $(X_1, X_2) = (f(U), g(U))$ for functions f and g which shows how (X_1, X_2) can be simulated.
b) With the help of a), use software and numerical integration to calculate the minimal correlation for two standard exponential random variables.

Exercise 7.13 (Maximal correlation for Pareto Type I random variables)
Compute the maximal correlation for a pair of random variables (X_1, X_2) with Pareto Type I marginal distribution functions of the form $F_j(x) = 1 - 1/x^{\theta_j}$, $x \geq 1$, $\theta_j > 2$, $j \in \{1,2\}$.

Exercise 7.14 (Copulas and correlation ordering)
Let (X_1, Y_1) have copula C_1 and (X_2, Y_2) have copula C_2. Assume moreover that $X_1 \overset{\mathrm{d}}{=} X_2$ and $Y_1 \overset{\mathrm{d}}{=} Y_2$ and that the marginal distributions of X_j and Y_j have finite second moments for $j \in \{1,2\}$. Use Höffding's Lemma to show that $C_1(\boldsymbol{u}) \leq C_2(\boldsymbol{u})$, $\boldsymbol{u} \in [0,1]^2$, implies that $\rho(X_1, Y_1) \leq \rho(X_2, Y_2)$.

Exercise 7.15 (Copula-based dependence measures under simple transformations)

a) Let (X_1, X_2) be a random vector with continuous margins. Compute $\rho_{\mathrm{S}}(-X_1, X_2^3)$ in terms of $\rho_{\mathrm{S}}(X_1, X_2)$.
b) Let $(U_1, U_2) \sim C$ for some copula C with upper tail dependence coefficient λ_u. What is the upper tail dependence coefficient of the copula of (U_1, U_2^2)?

Exercise 7.16 (Archimedean generator properties)
Let $\psi : [0, \infty) \to [0,1]$ be a generator of a d-dimensional Archimedean copula C.

a) Show that $\lim_{t\to\infty} \psi(t) = 0$ must hold.
b) Show that $\psi(0) = 1$ must hold.
c) Show that $\psi(ct)$, $t \geq 0$, generates C for all $c > 0$.
d) Show that the lower tail dependence coefficient of the bivariate copula satisfies $\lambda_{\mathrm{l}} = \lim_{t\to\infty} \frac{\psi(2t)}{\psi(t)}$.

Exercise 7.17 (Clayton copulas)

a) Show that the function $\psi(t) = (1+t)^{-1/\theta}$, $t \geq 0$, for $\theta > 0$ generates an Archimedean copula in all dimensions $d \geq 2$ and that the resulting copula is the Clayton copula

$$C_\theta(\boldsymbol{u}) = \Bigl(\sum_{j=1}^d u_j^{-\theta} - d + 1\Bigr)^{-1/\theta}, \quad \boldsymbol{u} \in [0,1]^d,\ \theta > 0.$$

b) Compute the lower tail dependence coefficient λ_{l} of the bivariate C_θ.

c) Show that C_θ converges to the independence copula as $\theta \to 0$.

d) Show that C_θ converges to the comonotonicity copula as $\theta \to \infty$.

Exercise 7.18 (Outer power Archimedean copula)

Let ψ be an Archimedean generator which generates an Archimedean copula C. The corresponding *outer power Archimedean copula* is the Archimedean copula generated by $\tilde{\psi}(t) = \psi(t^{1/\beta})$, $\beta \in [1, \infty)$. This provides a convenient method for introducing an extra parameter into an Archimedean copula and can give more flexibility for modelling tail dependencies in both upper and lower tails.

a) Show that Kendall's tau of the outer power Archimedean copula generated by $\tilde{\psi}$ is given by

$$\tilde{\rho}_\tau = 1 - (1 - \rho_\tau)/\beta,$$

where ρ_τ denotes Kendall's tau of the Archimedean copula C generated by ψ.

b) Show that the coefficients of lower and upper tail dependence of the outer power Archimedean copula generated by $\tilde{\psi}$ are given by

$$\tilde{\lambda}_{\mathrm{l}} = 2^{1/\beta} \lim_{t\to\infty} \frac{\psi'(2^{1/\beta}t)}{\psi'(t)} \quad \text{and} \quad \tilde{\lambda}_{\mathrm{u}} = 2 - 2^{1/\beta} \lim_{t\to 0} \frac{\psi'(2^{1/\beta}t)}{\psi'(t)},$$

respectively.

c) Let $\psi(t) = (1+t)^{-1/\theta}$, $\theta \in (0, \infty)$, denote the Clayton generator (with Kendall's tau $\rho_\tau = \theta/(\theta + 2)$). Compute Kendall's tau and the coefficients of tail dependence of the outer power Clayton copula generated by $\tilde{\psi}(t) = \psi(t^{1/\beta})$, $\beta \in [1, \infty)$.

Exercise 7.19 (Farlie–Gumbel–Morgenstern copula)

For $\theta \in [-1, 1]$, let

$$C(u_1, u_2) = u_1 u_2 + \theta(1 - u_1)(1 - u_2)u_1 u_2, \quad u_1, u_2 \in [0, 1].$$

a) Show that C is a copula.

b) Show that C is exchangeable.

c) Show that C is radially symmetric.

d) Compute Spearman's rho for $(U_1, U_2) \sim C$. Considering the range of ρ_{S}, what drawback does the Farlie–Gumbel–Morgenstern copula family have for statistical applications?

e) Compute the coefficient of lower tail dependence λ_{l}. Without further calculation, write down the coefficient of upper tail dependence λ_{u}.

Exercise 7.20 (Kendall's tau as a correlation)

The Kendall's tau of two continuously distributed random variables X_1, X_2 can be shown to be equal to

$$\rho_\tau(X_1, X_2) = 4\mathbb{P}(X_1 < X_1', X_2 < X_2') - 1,$$

where (X_1', X_2') denotes an independent copy of (X_1, X_2); see MFE (2015, equation (7.29)). Show that Kendall's tau can be expressed in terms of Pearson's correlation coefficient as

$$\rho_\tau(X_1, X_2) = \rho(I(X_1 \le X_1'), I(X_2 \le X_2')).$$

Exercise 7.21 (Blomqvist's beta)
Blomqvist's beta is a further copula-based measure of dependence. For random variables X_1 and X_2 with continuous distribution functions F_1 and F_2 it is defined by

$$\rho_\beta(X_1, X_2) = \mathbb{P}\big((X_1 - m_1)(X_2 - m_2) > 0\big) - \mathbb{P}\big((X_1 - m_1)(X_2 - m_2) < 0\big),$$

where $m_j = F_j^{\leftarrow}(1/2)$, $j \in \{1, 2\}$.

a) Show that $\rho_\beta(X_1, X_2)$ can be expressed in terms of the copula C of (X_1, X_2) alone.
b) Show that $\rho_\beta(X_1, X_2)$ can be expressed as

$$\rho_\beta(X_1, X_2) = \frac{C(1/2, 1/2) - \Pi(1/2, 1/2)}{M(1/2, 1/2) - \Pi(1/2, 1/2)},$$

where Π denotes the independence copula and M the comonotonicity copula.
c) Show that $\rho_\beta(X_1, X_2)$ can be expressed in terms of Pearson's correlation coefficient as

$$\rho_\beta(X_1, X_2) = \rho\big(I(X_1 \le m_1), I(X_2 \le m_2)\big).$$

Exercise 7.22 (Convex combinations of copulas)
Let C_1 and C_2 be two d-dimensional copulas and define, for $\gamma \in [0, 1]$,

$$C(\boldsymbol{u}) = \gamma C_1(\boldsymbol{u}) + (1 - \gamma) C_2(\boldsymbol{u}), \quad \boldsymbol{u} \in [0, 1]^d.$$

Taking convex combinations in this manner provides a further way of expanding the set of copulas available to risk modellers.

a) Show that C is a copula by verifying the three characterizing properties of a copula in MFE (2015, Section 7.1.1).
b) Suppose that we can generate random vectors $\boldsymbol{U}_1 \sim C_1$ and $\boldsymbol{U}_2 \sim C_2$. By recognizing that C is a mixture model, write down an algorithm for generating a random vector $\boldsymbol{U} \sim C$ and verify its correctness.
c) For $d = 2$, show that the Spearman's rho ρ_{S} of C is given by $\rho_{\mathrm{S}} = \gamma \rho_{\mathrm{S},1} + (1 - \gamma)\rho_{\mathrm{S},2}$, where $\rho_{\mathrm{S},k}$ denotes the Spearman's rho of C_k, $k \in \{1, 2\}$.
d) For $d = 2$, show that the Kendall's tau ρ_τ of C is given by

$$\rho_\tau = \gamma^2 \rho_{\tau,1} + 2\gamma(1 - \gamma)\big(4\mathbb{E}(C_1(\boldsymbol{U}_2)) - 1\big) + (1 - \gamma)^2 \rho_{\tau,2},$$

where $\rho_{\tau,k}$ denotes the Kendall's tau C_k, $k \in \{1, 2\}$.

Hint. You may use the fact that $\mathbb{E}(C_2(\boldsymbol{U}_1)) = \mathbb{E}(C_1(\boldsymbol{U}_2))$.

Exercise 7.23 (Bivariate Marshall–Olkin copulas)
Common shock models have applications in insurance modelling and credit risk modelling. In the latter context consider three firms and for $j \in \{1,2,3\}$ let T_j denote the time to bankruptcy for firm j due to reasons that are exclusive to each firm (for example bad management). Assume moreover that firms 1 and 2 are suppliers to firm 3 and will both fail immediately if firm 3 fails first.

Let $T_j \sim \mathrm{Exp}(\lambda_j)$, $j \in \{1,2,3\}$ and assume that T_1, T_2 and T_3 are independent. The actual times to bankruptcy for firms 1 and 2 are given by

$$X_j = \min\{T_j, T_3\}, \quad j \in \{1,2\}.$$

The survival copula of (X_1, X_2) is a *Marshall–Olkin copula*; see Marshall and Olkin (1967).

a) Show that the survival copula of (X_1, X_2) is given by

$$C(u_1, u_2) = \min\{u_1 u_2^{1-\alpha_2}, u_1^{1-\alpha_1} u_2\}, \quad u_1, u_2 \in [0,1],$$

where $\alpha_j = \lambda_3/(\lambda_j + \lambda_3)$, $j \in \{1,2\}$; note that one typically includes the boundary cases $\alpha_j \in \{0,1\}$, $j \in \{1,2\}$, in this representation.

b) Let $(U_1, U_2) \sim C$. Find a stochastic representation of the form $(U_1, U_2) = (f_1(V_1, V_3), f_2(V_2, V_3))$ for $V_1, V_2, V_3 \overset{\text{ind.}}{\sim} \mathrm{U}(0,1)$ and functions f_1 and f_2 that you should determine.

c) Compute the tail dependence coefficients $\lambda_{\mathrm{l}}, \lambda_{\mathrm{u}}$ for $(U_1, U_2) \sim C$ and any $\alpha_1, \alpha_2 \in [0,1]$.

Exercise 7.24 (Sampling copulas and meta-distributions)
Use statistical software to reproduce and extend the analyses in MFE (2015, Example 7.13). In other words consider the following four bivariate copulas:

i) the Gauss copula with parameter $\rho = 0.7$;
ii) the Gumbel copula with parameter $\theta = 2$;
iii) the Clayton copula with parameter $\theta = 2.2$; and
iv) the t copula with parameters $\nu = 4$ and $\rho = 0.71$.

a) Generate random samples of size 2000 from the four copulas and produce scatterplots.

b) Generate and plot samples from meta-distributions based on the four copulas above combined with standard normal margins.

c) The copula parameters were chosen to give meta-distributions with a linear correlation of approximately 70%. Now try to generate samples from meta-distributions which combine the four copulas above with unit exponential margins. Change the copula parameters for the Gumbel, Clayton and t copulas so that they all have the same value for Kendall's tau as the Gauss copula.

The R package `copula` can be used to answer this question.

Exercise 7.25 (Fitting copulas to equity return data)
Take the same data that were analysed in Exercise 6.23, in other words daily log-return data from 2005 to 2012 for the following 10 components of the Dow Jones index: Apple, Cisco, Disney, IBM, Intel, McDonald's, Microsoft, Nike, Proctor & Gamble, and Walmart.

a) Fit both the Gauss copula and the t copula to the log-returns of these data using the pseudo maximum-likelihood approach to take care of the unknown marginal distributions. What are the fitted copula parameters?
b) Compare the Akaike information criteria (AIC numbers) for the two models and confirm that the t copula is clearly the superior model for the dependence structure.
c) Repeat the analysis for monthly log-return data.

The R package `copula` contains fitting functions for these copulas.

Advanced

Exercise 7.26 (The conditional copula of a Clayton copula)
Let (U_1, U_2, U_3) be distributed according to a Clayton copula with parameter $\theta > 0$. For $u_3 \in (0,1)$, derive the copula of $(U_1, U_2) \mid U_3 = u_3$, that is the conditional copula of (U_1, U_2) given that $U_3 = u_3$.

Exercise 7.27 (A flexible bivariate copula construction)
In this exercise we study a flexible copula construction that can be used to build copulas with a variety of different properties.

Let $g_1, g_2 : [0,1] \to [0,1]$ be continuously differentiable functions satisfying $g_j(0) = g_j(1) = 0$, $j \in \{1,2\}$, and $g_1'(u_1)g_2'(u_2) \geq -1$ for all $u_1, u_2 \in [0,1]$. Consider

$$C(u_1, u_2) = u_1 u_2 + g_1(u_1)g_2(u_2), \quad u_1, u_2 \in [0,1].$$

a) Show that C defines a copula.
b) Compute $\rho = \rho(U_1, U_2) = \rho_S(U_1, U_2)$ for $(U_1, U_2) \sim C$.
c) Show that $\rho = 0$ if g_1 or g_2 are point symmetric about $1/2$ (meaning that $g_i(u) = -g_i(1-u)$).
d) Hence construct a parametric copula family satisfying $\rho = 0$ for all its members.
e) Show that C is radially symmetric if g_1 and g_2 are symmetric about $1/2$.
f) Hence construct a copula that is radially symmetric but non-exchangeable.

Exercise 7.28 (The Liebscher construction)
Let C_1 and C_2 be two copulas and $(U_1, V_1) \sim C_1$ and $(U_2, V_2) \sim C_2$ two independent pairs of random variables. For $\alpha, \beta \in (0,1)$, define

$$\begin{aligned} X_1 &= \max\{U_1^{1/\alpha},\ U_2^{1/(1-\alpha)}\}, \\ X_2 &= \max\{V_1^{1/\beta},\ V_2^{1/(1-\beta)}\}. \end{aligned}$$

a) Show that the copula of (X_1, X_2) is given by

$$C(u_1, u_2) = C_1(u_1^{\alpha}, u_2^{\beta})C_2(u_1^{1-\alpha}, u_2^{1-\beta}), \quad u_1, u_2 \in [0,1].$$

b) Compute the four limiting copulas as (α, β) tends to $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$.

c) For $\alpha = \beta$, compute the coefficient of lower tail dependence λ_{l} of C in terms of the coefficient of lower tail dependence $\lambda_{\mathrm{l},j}$ of C_j, $j \in \{1, 2\}$.

Exercise 7.29 (The Rosenblatt transformation and the conditional sampling method)
For a d-dimensional copula C, $\boldsymbol{U} \sim C$, $j \in \{2, \dots, d\}$ and $u_1, \dots, u_{j-1} \in (0, 1)$, one can define the conditional distribution

$$C_{j|1,\dots,j-1}(u_j \,|\, u_1, \dots, u_{j-1}) = \mathbb{P}(U_j \le u_j \,|\, U_1 = u_1, \dots, U_{j-1} = u_{j-1}), \quad u_j \in [0, 1],$$

of C. It can be shown that if C admits continuous partial derivatives with respect to the first $j-1$ arguments, then

$$C_{j|1,\dots,j-1}(u_j \,|\, u_1, \dots, u_{j-1}) = \frac{\frac{\partial^{j-1}}{\partial u_{j-1} \dots \partial u_1} C^{(1,\dots,j)}(u_1, \dots, u_j)}{\frac{\partial^{j-1}}{\partial u_{j-1} \dots \partial u_1} C^{(1,\dots,j-1)}(u_1, \dots, u_{j-1})}, \tag{E1}$$

where $C^{(1,\dots,j)}(u_1, \dots, u_j) = C(u_1, \dots, u_j, 1, \dots, 1)$ denotes the marginal copula of C corresponding to the first j arguments (with the convention that $C^{(1)}(u_1) = u_1$).

The transformation $R_C : (0,1)^d \to (0,1)^d$ defined by $R_C(\boldsymbol{u}) = \boldsymbol{u}'$ with

$$\begin{aligned} u_1' &= u_1, \\ u_2' &= C_{2|1}(u_2 \,|\, u_1), \\ &\;\;\vdots \\ u_d' &= C_{d|1,\dots,d-1}(u_d \,|\, u_1, \dots, u_{d-1}); \end{aligned}$$

is known as the *Rosenblatt transformation* (Rosenblatt (1952)). If $\boldsymbol{U} \sim C$, it can be shown that $R_C(\boldsymbol{U})$ follows the independence copula Π and this fact is used in goodness-of-fit tests; see Genest et al. (2009).

Consider the corresponding inverse transformation $R_C^{\leftarrow} : (0,1)^d \to (0,1)^d$ defined by $R_C^{\leftarrow}(\boldsymbol{u}') = \boldsymbol{u}$ with

$$\begin{aligned} u_1 &= u_1', \\ u_2 &= C_{2|1}^{\leftarrow}(u_2' \,|\, u_1), \\ &\;\;\vdots \\ u_d &= C_{d|1,\dots,d-1}^{\leftarrow}(u_d' \,|\, u_1, \dots, u_{d-1}), \end{aligned}$$

where $C_{j|1,\dots,j-1}^{\leftarrow}$ denotes the quantile function of $u_j \mapsto C_{j|1,\dots,j-1}(u_j \,|\, u_1, \dots, u_{j-1})$. One can show that if $\boldsymbol{U}' \sim \Pi$, then $R_C^{\leftarrow}(\boldsymbol{U}') \sim C$. This inverse Rosenblatt transformation can be used for sampling from C and the corresponding algorithm is known as the *conditional distribution method*.

a) Compute $C_{j|1,\dots,j-1}$ and $C_{j|1,\dots,j-1}^{\leftarrow}$ for an Archimedean copula with d-times continuously differentiable generator ψ satisfying $\psi(t) > 0$ for all $t \in [0, \infty)$.

b) Compute $C_{j|1,\dots,j-1}$ and $C^{\leftarrow}_{j|1,\dots,j-1}$ for the Clayton copula.

Exercise 7.30 (Archimedean Marshall–Olkin copulas)
Let $(W_1, W_2, W_3) \sim C_0$ where C_0 is an Archimedean copula with generator $\psi_0(t)$. Let

$$X_j = \min\{T_j, T_{12}\}, \quad j \in \{1, 2\},$$

where $(T_1, T_2, T_{12}) = (-\log W_1, -\log W_2, -\log W_3)$. Show that the survival copula C of (X_1, X_2) is given by

$$C(u_1, u_2) = C_0(G_0(u_1), G_0(u_2), G_0(\min\{u_1, u_2\})), \quad u_1, u_2 \in [0, 1],$$

where $G_0(u) = \psi_0(\psi_0^{-1}(u)/2)$, $u \in [0, 1]$.

Exercise 7.31 (Axioms of concordance)
Spearman's rho and Kendall's tau are measures of concordance. According to Scarsini (1984), such measures should fulfill a set of desirable axioms. Let (X, Y) be a pair of random variables with continuous marginal distribution functions and joint distribution function F. Let the measure of concordance be $\kappa(F) = \kappa(X, Y)$. The axioms are:

i) (Domain) $\kappa(X, Y)$ is defined for every such pairs (X, Y);
ii) (Range) $-1 \le \kappa \le 1$ with $\kappa(X, Y) = -1$ for countermonotonic random variables and $\kappa(X, Y) = 1$ for comonotonic random variables;
iii) (Symmetry) $\kappa(X, Y) = \kappa(Y, X)$;
iv) (Independence) If X and Y are independent, then $\kappa(X, Y) = 0$;
v) (Change of sign) $\kappa(-X, Y) = -\kappa(X, Y)$;
vi) (Coherence) If (X_1, Y_1) and (X_2, Y_2) are pairs of random variables with copulas $C_1 \le C_2$ ordered pointwise, then $\kappa(X_1, Y_1) \le \kappa(X_2, Y_2)$;
vii) (Continuity) If $(F_n)_{n\in\mathbb{N}}$ is a sequence of continuous bivariate distribution functions which converges pointwise to the continuous distribution function F, then $\lim_{n\to\infty} \kappa(F_n) = \kappa(F)$.

a) Verify these axioms for Spearman's rho.
b) Provide an example which shows that $\kappa(X, Y) = 0$ does not imply independence of X, Y.

   Hint. Consider the uniform distribution on the unit circle.

8 Aggregate Risk

Review

Exercise 8.1 (Acceptance sets)

Let ϱ be a monotone and translation-invariant risk measure.

a) Provide a mathematical definition of the acceptance set A_ϱ corresponding to ϱ.

b) How are convexity and coherence of ϱ related to the geometric properties of A_ϱ.

c) Describe A_ϱ for $\varrho = \mathrm{VaR}_\alpha$, $\alpha \in (0, 1)$.

Exercise 8.2 (Dual representation of ES)

a) State the dual representation of expected shortfall ES_α, $\alpha \in [0, 1)$.

b) By definition, ES_α is increasing in α. How can this be seen from the dual representation?

c) Use the dual representation to show that ES_α is subadditive.

Exercise 8.3 (Distortion)

a) Give the definition of a distortion risk measure.

b) Explain how distortion risk measures can be represented in terms of expected shortfall.

c) Use this representation to show that distortion risk measures are comonotone additive and coherent.

Exercise 8.4 (Subadditivity of VaR for jointly normal risks)

Suppose $\boldsymbol{X} = (X_1, \dots, X_d) \sim \mathrm{N}_d(\boldsymbol{\mu}, \Sigma)$ and let $\mathcal{M}$ denote the space of all risks of the form $L = \boldsymbol{\lambda}'\boldsymbol{X} = \sum_{j=1}^d \lambda_j X_j$. Show by calculation that VaR_α is subadditive on $\mathcal{M}$ for all $\alpha \in [1/2, 1)$.

Exercise 8.5 (Maximal ES of the sum)

Let $L_1 \sim F_1$ and $L_2 \sim F_2$ with finite second moments. Explain whether or not $\mathrm{ES}_\alpha(L_1 + L_2)$ is maximized when the correlation between L_1 and L_2 is maximal.

Exercise 8.6 (Revisiting Markowitz)

Consider linear portfolios in the set

$$\mathcal{M} = \left\{L : L = \boldsymbol{\lambda}'\boldsymbol{X},\ \boldsymbol{\lambda} \in \mathbb{R}^d\right\}, \tag{E1}$$

where $\boldsymbol{X}$ is a fixed d-dimensional random vector of risk factors defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and where $\boldsymbol{\lambda}$ represents portfolio weights. Denote by $L(\boldsymbol{\lambda}) = \sum_{j=1}^d \lambda_j X_j$ the loss generated by the portfolio weights $\boldsymbol{\lambda}$ and assume that $\boldsymbol{X}$ has an elliptical distribution with finite variance. Fix some $\alpha > 0.5$ and suppose that VaR_α is used as a risk measure.

Consider the problem of finding the VaR_α-minimizing linear portfolio in the set of all portfolios whose expected loss is equal to a given constant c, that is consider the problem

$$\min_{\boldsymbol{\lambda}\in\mathbb{R}^d} \mathrm{VaR}_\alpha(L(\boldsymbol{\lambda})) \quad \text{such that} \quad \mathbb{E}(L(\boldsymbol{\lambda})) = c.$$

Explain why the solution $\boldsymbol{\lambda}^*$ of this problem is identical to the classical Markowitz portfolio, that is to the solution of the problem $\min_{\boldsymbol{\lambda}\in\mathbb{R}^d} \mathrm{var}(L(\boldsymbol{\lambda}))$ such that $\mathbb{E}(L(\boldsymbol{\lambda})) = c$.

Exercise 8.7 (On correlation adjusted summation in risk aggregation)
Let $\mathrm{EC}_1, \dots, \mathrm{EC}_d$ denote (economic) capital amounts for d business lines of a financial firm with associated losses $L_1, \dots, L_d$ and let $L = \sum_{j=1}^d L_j$ denote the total loss.

a) Discuss strengths and weaknesses of the simple summation formula

$$\mathrm{EC} = \sum_{j=1}^d \mathrm{EC}_j$$

and of the correlation adjusted summation formula

$$\mathrm{EC} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d \rho_{ij}\, \mathrm{EC}_i\, \mathrm{EC}_j},$$

for risk aggregation; here $\rho_{ij} \in [0,1]$, $1 \le i, j \le d$, are parameters with $\rho_{ii} = 1$ and $\rho_{ij} = \rho_{ji}$ that are usually referred to as 'correlations'.

b) If risk is measured with $\varrho = \mathrm{ES}_\alpha$, is simple summation always conservative in the sense of providing an upper bound for $\varrho(L)$? What about $\varrho = \mathrm{VaR}_\alpha$?

c) Take $d = 2$ and suppose that L_j is lognormally distributed with parameters μ_j and σ_j, $j \in \{1, 2\}$. Suppose that the firm decides to use correlation adjusted summation with some parameter $\rho = \rho_{12} \in (0, 1]$. Does the interpretation of ρ as correlation of L_1 and L_2 always make sense?

Exercise 8.8 (Economic properties of a capital allocation principle)
Consider an insurance company with d business lines, producing random losses $L_1, \dots, L_d$, respectively, so that the total loss is $L = \sum_{j=1}^d L_j$. Let ϱ be a positive-homogenous risk measure and let $\varrho(L)$ be the risk capital for the entire company. In this context, a capital allocation principle allocates the economic capital amounts $\mathrm{AC}_1, \dots, \mathrm{AC}_d$ to the respective business lines.

a) Explain briefly why capital allocation principles are used in risk adjusted performance measurement.

b) Give the definition of the Euler capital allocation principle in the above context and explain why in the above context RORAC-compatibility and the existence of a diversification benefit (two properties of the Euler capital allocation principle) are desirable properties of a capital allocation principle.

Basic

Exercise 8.9 (Acceptance sets for generalized scenarios)
Let $\Omega = \{\omega_1, \omega_2\}$ and let $\mathbb{Q}_1, \mathbb{Q}_2$ be probability measures on Ω such that $\mathbb{Q}_1(\{\omega_1\}) = 0.6$ and $\mathbb{Q}_2(\{\omega_1\}) = 0.3$. Consider the generalized scenario risk measure $\varrho(L) = \max\{\mathbb{E}^{\mathbb{Q}_1}(L),\ \mathbb{E}^{\mathbb{Q}_2}(L)\}$. By identifying a random variable L with a vector $\boldsymbol{l} = (L(\omega_1), L(\omega_2))$ we can identify the acceptance set of $\varrho(L)$ with a subset of $\mathbb{R}^2$. Describe this set.

Exercise 8.10 (Convexity of coherent risk measures)

a) Show that every coherent risk measure $\varrho : \mathcal{M} \to \mathbb{R}$ is a convex risk measure, that is ϱ satisfies

$$\varrho(\lambda L_1 + (1-\lambda)L_2) \le \lambda\varrho(L_1) + (1-\lambda)\varrho(L_2), \quad \lambda \in [0,1],\ L_1, L_2 \in \mathcal{M}.$$

b) Show that for positive-homogeneous risk measures convexity and coherence are equivalent.

c) Show by means of a counterexample that without positive homogeneity the statement in b) is false.

Exercise 8.11 (Standard deviation principle as a risk measure)
For a random variable L with finite second moment, suppose that we define the risk measure $\varrho(L) = \mathbb{E}(L) + k\,\mathrm{sd}(L)$ where k is a positive constant and sd denotes standard deviation. This risk measure is known as standard deviation principle. Show that this risk measure is translation invariant, subadditive and positive-homogeneous, but not monotone in general.

Hint. For a counterexample for monotonicity let $L_1 = 0$ and suppose that, for sufficiently large n,

$$L_2 = \begin{cases} 0 & \text{with probability } 1 - 1/n, \\ -1 & \text{with probability } 1/n. \end{cases}$$

Exercise 8.12 (Standardized and strictly convex risk measures)
Consider a strictly convex risk measure ϱ that is standardized so that $\varrho(0) = 0$. Show that for $\lambda > 1$ it holds that $\varrho(\lambda L) > \lambda\varrho(L)$ and give an economic interpretation of this inequality.

Hint. A convex risk measure ϱ is strictly convex if for L_1 and L_2 in the domain of ϱ with $L_1 \ne L_2$ and $\lambda \in (0,1)$ one has $\varrho\big(\lambda L_1 + (1-\lambda)L_2\big) < \lambda\varrho(L_1) + (1-\lambda)\varrho(L_2)$.

Exercise 8.13 (Monotonicity of expectiles in the confidence level)
Show that the α-expectile $e_\alpha(L)$ of a loss L with $\mathbb{E}(|L|) < \infty$ is increasing in the confidence level α.

Exercise 8.14 (Expectiles of the standard uniform distribution)

a) Show that the α-expectile $e_\alpha(L)$ of a loss L with $\mathbb{E}(|L|) < \infty$ satisfies

$$e_\alpha(L) = \mathbb{E}(L) + \frac{2\alpha - 1}{1-\alpha}\mathbb{E}((L-y)^+) = \mathbb{E}(L) + \frac{2\alpha-1}{\alpha}\mathbb{E}((L-y)^-),$$

where $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$.

b) Now assume that $L \sim \mathrm{U}(0,1)$.
   i) Compute $e_\alpha(L)$ using a).
   ii) Compute $e_\alpha(L)$ using MFE (2015, Proposition 8.23).
c) What is $e_\alpha(L)$ in the case where $L \sim \mathrm{U}(a,b)$ for $a < b$.

Exercise 8.15 (Expectiles for exponential and Pareto risks)

a) Compute $e_\alpha(L)$ for $L \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$.

   Hint. You may use the Lambert W function, defined as the inverse of the function $x \mapsto x\exp(x)$.
b) Compute $e_\alpha(L)$ for $L \sim \mathrm{Pa}(2,\kappa)$, $\kappa > 0$.

Exercise 8.16 (Stress test risk measure)

Consider a set of potential losses of a financial institution which are of the form $L = l(\boldsymbol{x})$ where $\boldsymbol{x}$ denotes a d-dimensional vector of risk-factor changes and where l is the risk mapping associated with the loss L. Fix a set $S \subseteq \mathbb{R}^d$ of possible risk-factor changes (the scenarios). A stress test risk measure is then given by

$$\varrho_S(L) = \sup\{l(\boldsymbol{x}) \,:\, \boldsymbol{x} \in S\},$$

that is the worst loss over all scenarios which belong to S. Show that $\varrho_S(L)$ is coherent.

Exercise 8.17 (Euler principle for multivariate normal risks)

Assume that $(L_1,\dots,L_d) \sim \mathrm{N}_d(\mathbf{0},\Sigma)$. For the risk measure $\varrho = \mathrm{ES}_\alpha$ compute the Euler capital allocations AC_j, the so-called expected shortfall contributions.

Hint. You may use the fact that $\mathrm{ES}_\alpha(L) = \mu + \sigma\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$ for $L \sim \mathrm{N}(\mu,\sigma^2)$.

Advanced

Exercise 8.18 (Dual representation of convex risk measures)

This exercise is concerned with the computations in MFE (2015, Example 8.12) which illustrates the proof of the existence of a dual representation for a convex risk measure. Consider $(\Omega,\mathcal{F},\mathbb{P})$ with $\Omega = \{\omega_1,\omega_2\}$, $\mathcal{F}$ the power set of Ω and $\mathbb{P}(\{\omega_k\}) = 0.5$ for $k \in \{1,2\}$. Consider the convex risk measure $\varrho(L) = \log\mathbb{E}(e^L)$. For a loss L satisfying $\varrho(L) = 0$ construct the probability measure $\mathbb{Q}_L$ in MFE (2015, Theorem 8.11, Step 2).

Hint. Show that $\mathcal{A}_\varrho = \{(\ell_1,\ell_2) : \ell_2 \le \log(2 - e^{\ell_1})\}$.

Exercise 8.19 (Representation of ES)

For an integrable loss L with distribution function F_L expected shortfall may be written as

$$\mathrm{ES}_\alpha(L) = \frac{\mathbb{E}(LI_{\{L>F_L^{\leftarrow}(\alpha)\}}) + F_L^{\leftarrow}(\alpha)\big(1-\alpha-\bar{F}_L(F_L^{\leftarrow}(\alpha))\big)}{1-\alpha}; \tag{E1}$$

see MFE (2015, Proposition 8.13). In this exercise we give several related representations for ES_α.

a) Define the *tail value-at-risk (*TVaR*)* (also known as *conditional tail expectation (*CTE*)*) at confidence level $\alpha \in (0,1)$ by

$$\mathrm{TVaR}_\alpha(L) = \mathbb{E}(L \,|\, L > \mathrm{VaR}_\alpha(L)).$$

Show that

$$\mathrm{ES}_\alpha(L) = \mathrm{TVaR}_\alpha(L)\frac{\bar{F}_L(F_L^{\leftarrow}(\alpha))}{1-\alpha} + \mathrm{VaR}_\alpha(L)\Bigl(1 - \frac{\bar{F}_L(F_L^{\leftarrow}(\alpha))}{1-\alpha}\Bigr)$$

and interpret this formula.

b) Let

$$I^{(\alpha)}_{\{L>q\}} = \begin{cases} I_{\{L>q\}}, & \mathbb{P}(L=q)=0, \\ I_{\{L>q\}} + \frac{1-\alpha-\bar{F}_L(q)}{\mathbb{P}(L=q)} I_{\{L=q\}}, & \mathbb{P}(L=q)>0. \end{cases}$$

Show that ES_α allows for the representation

$$\mathrm{ES}_\alpha(L) = \frac{\mathbb{E}(L I^{(\alpha)}_{\{L>F_L^{\leftarrow}(\alpha)\}})}{1-\alpha}.$$

Exercise 8.20 (CoVaR and systemic risk)

Let L_1 and L_2 be two continuously distributed losses interpreted as either losses of two firms in a financial system or of one firm and the financial system as a whole. The *conditional value-at-risk* CoVaR of L_2 given L_1 at confidence level $\alpha \in (0,1)$ is a risk measure defined by

$$\mathrm{CoVaR}_{\alpha,\beta}(L_2 \,|\, L_1) = F^{\leftarrow}_{L_2|L_1 \ge \mathrm{VaR}_\beta(L_1)}(\alpha),$$

where $F_{L_2|L_1\ge \mathrm{VaR}_\beta(L_1)}(x) = \mathbb{P}(L_2 \le x \,|\, L_1 \ge \mathrm{VaR}_\beta(L_1))$. The *systemic risk contribution* of L_1 to L_2 at confidence level $\alpha \in (0,1)$ is defined by

$$\Delta\,\mathrm{CoVaR}_\alpha(L_2 \,|\, L_1) = \mathrm{CoVaR}_{\alpha,\alpha}(L_2 \,|\, L_1) - \mathrm{CoVaR}_{\alpha,0.5}(L_2 \,|\, L_1).$$

a) Interpret $\mathrm{CoVaR}_{\alpha,\beta}(L_2 \,|\, L_1)$ and $\Delta\,\mathrm{CoVaR}_\alpha(L_2 \,|\, L_1)$.

b) If $(L_1, L_2) \sim F$ with margins F_1, F_2 and copula C, find an expression for $F_{L_2|L_1\ge \mathrm{VaR}_\beta(L_1)}$.

c) Derive an expression for $\mathrm{CoVaR}_{\alpha,\beta}(L_2 \,|\, L_1)$ for $C(\boldsymbol{u}) = \gamma M(\boldsymbol{u}) + (1-\gamma)\Pi(\boldsymbol{u})$, $\boldsymbol{u} \in [0,1]^2$, where M and Π denote the comonotonicity and the independence copulas and $\gamma \in [0,1]$. Comment on the special cases $\gamma \in \{0,1\}$.

d) Under the setup of c) and for $L_2 \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$, compute the systemic risk contribution $\Delta\,\mathrm{CoVaR}_\alpha(L_2 \,|\, L_1)$, $\alpha \in (1/2, 1)$. Comment on the special cases $\gamma \in \{0,1\}$.

Note. CoVaR and Δ CoVaR were introduced by Adrian and Brunnermeier (2016) but with conditioning event $L_1 = \mathrm{VaR}_\beta(L_1)$ instead of $L_1 \ge \mathrm{VaR}_\beta(L_1)$. The former conditioning event can also be found in the definition of CoVaR in other literature. Mainik and Schaanning (2014) give reasons why the conditioning event $L_1 \ge \mathrm{VaR}_\beta(L_1)$ is to be preferred.

Exercise 8.21 (Superadditivity of VaR under worst-case dependence)
For $\alpha \in (0,1)$ and $L_1 \sim \mathrm{U}(0,1)$ let L_2 be almost surely given by

$$L_2 = \begin{cases} L_1, & L_1 < \alpha, \\ 1+\alpha-L_1, & L_1 \geq \alpha. \end{cases}$$

a) Show that $L_2 \sim \mathrm{U}(0,1)$.
b) Let $\alpha \in (0,1)$ and $\varepsilon \in (0,(1-\alpha)/2)$. Compute $\mathrm{VaR}_{\alpha+\varepsilon}(L_1+L_2)$ and compare it to $\mathrm{VaR}_{\alpha+\varepsilon}(L_1)+\mathrm{VaR}_{\alpha+\varepsilon}(L_2)$.
c) For $\varepsilon \in (0,(1-\alpha)/2)$ compute the upper bound $\overline{\mathrm{VaR}}_{\alpha+\varepsilon}(\mathcal{S}_2)$ from MFE (2015, Proposition 8.31) and relate this to the result of b).

Exercise 8.22 (A proof of subadditivity of ES based on Höffding's Lemma)
Let L be a random variable with $\mathbb{E}(|L|) < \infty$ and recall that L admits the stochastic representation $L = F_L^{\leftarrow}(U)$ where $U \sim \mathrm{U}(0,1)$ and F_L is the distribution function of L.

a) By considering $Y = I_{\{U \geq \alpha\}}$ for $\alpha \in (0,1)$, show the validity of the representation

$$\mathrm{ES}_\alpha(L) = \frac{1}{1-\alpha}\sup\bigl\{\mathbb{E}(L\tilde{Y}) \,:\, \tilde{Y} \sim \mathrm{B}(1,1-\alpha)\bigr\},$$

where the supremum is understood as being taken over all joint distributions of $(L,\tilde{Y})$ such that L has marginal distribution function F_L and $\tilde{Y} \sim \mathrm{B}(1,1-\alpha)$.
b) Conclude that expected shortfall is subadditive.

Exercise 8.23 (A basic version of the rearrangement algorithm)
Let L_1, L_2 and L_3 be losses from three business lines of a financial firm with continuous distribution functions F_1, F_2 and F_3 respectively. Assume the firm measures risk with $\mathrm{VaR}_\alpha(L)$ based on the total loss $L = L_1 + L_2 + L_3$ for some confidence level $\alpha \in (0,1)$. The following questions address elementary steps in understanding the rearrangement algorithm of Embrechts et al. (2013) for computing lower and upper bounds on the worst value-at-risk $\overline{\mathrm{VaR}}_\alpha(L)$, that is the largest $\mathrm{VaR}_\alpha(L)$ under F_1, F_2, F_3. For simplicity, the basic version below describes the necessary steps to obtain one such bound on $\overline{\mathrm{VaR}}_\alpha(L)$ and we take it as an estimate for $\overline{\mathrm{VaR}}_\alpha(L)$.

a) Suppose the matrix

$$A = \begin{pmatrix} 2 & 3 & 8 \\ 4 & 1 & 1 \\ 1 & 5 & 2 \\ 3 & 7 & 4 \end{pmatrix}$$

contains in column $j \in \{1,2,3\}$ the largest $(1-\alpha)100\%$ of the losses incurred in business line j. The vector of row sums $(13,6,8,14)$ of A can thus be viewed as realizations of the largest $(1-\alpha)100\%$ of the total loss L and the minimal number, 6, is thus an approximate

$\mathrm{VaR}_\alpha(L)$ estimate, denoted by $\widehat{\mathrm{VaR}}_\alpha(L)$. By rearranging the losses in each column of A, we can mimic what happens to $\widehat{\mathrm{VaR}}_\alpha(L)$ under different types of copulas C of (L_1, L_2, L_3) in the joint tail.

Rearrange the losses in each column of A such that the rows of the rearranged matrix A mimic realizations obtained under the comonotonicity copula $C = M$. What is the corresponding $\mathrm{VaR}_\alpha(L)$ estimate $\widehat{\mathrm{VaR}}_\alpha(L)$? How would you need to rearrange each column of A if C equals the independence copula Π?

b) An estimate $\widehat{\overline{\mathrm{VaR}}}_\alpha(L)$ of worst value-at-risk $\overline{\mathrm{VaR}}_\alpha(L)$ over all copulas C of (L_1, L_2, L_3) in the joint tail can be constructed by rearranging the entries in each column of A such that the minimum of the row sums of the rearranged matrix becomes maximal (see iii) below for an explanation). To rearrange A, iterate over the columns of A and oppositely order the current column with respect to the sum over all other columns; note that two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are oppositely ordered if the smallest element of $\boldsymbol{a}$ lies next to the largest element of $\boldsymbol{b}$, the second-smallest next to the second-largest and so on. Assume this basic version of the so-called *rearrangement algorithm (RA)* is terminated if every column is oppositely ordered with respect to the sum over all other columns.

i) Apply this basic version of the RA to A for computing the estimate $\widehat{\overline{\mathrm{VaR}}}_\alpha(L)$ of worst value-at-risk $\overline{\mathrm{VaR}}_\alpha(L)$.

Hint. When rearranging column j, it is possible that the vector of row sums over all other columns contains ties, that is equal values. In this case, you are free to choose an order when rearranging the corresponding entries in column j.

ii) Compare your result with the case $C = M$ from a).

iii) Explain intuitively why each step of the RA can increase $\mathrm{VaR}_\alpha(L)$ and thus provide a way to approximate $\overline{\mathrm{VaR}}_\alpha(L)$.

Hint. Consider the case where all the losses are continuously distributed with finite first moment and suppose it is possible to rearrange the columns of A such that all row sums are (nearly) equal.

c) Apply the basic RA as described in b) to the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$$

for computing the estimate $\widehat{\overline{\mathrm{VaR}}}_\alpha(L)$ of $\overline{\mathrm{VaR}}_\alpha(L)$. Is the $\widehat{\overline{\mathrm{VaR}}}_\alpha(L)$ obtained this way indeed the largest possible $\overline{\mathrm{VaR}}_\alpha(L)$ estimate in this case?

9 Market Risk

Review

Exercise 9.1 (Mapping and loss operator)
Briefly explain in words what is meant by the mapping of a portfolio and explain the concept of a loss operator.

Exercise 9.2 (Unconditional and conditional approaches to market risk measurement)
Briefly explain the ideas behind the unconditional and conditional approaches to market risk measurement and summarize the pros and cons of each approach.

Exercise 9.3 (Comparison of variance–covariance, historical simulation and Monte Carlo)
List some advantages and disadvantages of the variance–covariance, historical simulation and Monte Carlo methods for measuring market risk.

Exercise 9.4 (Historical simulation)

a) Give a theoretical justification for the basic historical simulation method without volatility estimation.
b) Explain why this is an unconditional approach.
c) Sketch a method of extending historical simulation to a time-series context where volatility is changing.

Exercise 9.5 (Elicitability and its relevance to model validation)

a) Briefly explain the notion of elicitability for a law-invariant risk measure ϱ.
b) Explain why elicitability of a risk measure ϱ is useful for comparing risk measure forecasts.
c) Categorize the following risk measures according to whether they are elicitable or not: value-at-risk, expected shortfall, the expectile risk measure.

Exercise 9.6 (VaR violations)

a) Explain the concept of a violation of value-at-risk at level α.
b) What would be the properties of the time series of VaR_α violations produced by an ideal forecaster who always uses the correct conditional loss distribution to forecast losses?

Exercise 9.7 (Factor models for bond portfolios)
Describe the main differences between the Nelson-Siegel and PCA methods of building factor models for measuring the risk in a bond portfolio.

Basic

Exercise 9.8 (Mapping an interest-rate swap)

A (forward-start) interest-rate swap (IRS) is a contract with future start date T_0 where two parties agree to exchange payment streams at a set of times $T_1 < \cdots < T_n$ with $T_1 > T_0$. One party receives the amount $N\delta_i K$ at time T_i where N is the nominal value of the contract, K is a fixed interest rate and δ_i is the length of the interval $[T_{i-1}, T_i]$ expressed as a fraction of a year; in return the other party makes the floating payment $N\delta_i L(T_{i-1}, T_i)$ where $L(T_{i-1}, T_i)$ is the simply compounded spot interest rate at time T_{i-1} for the maturity time T_i (this is typically the LIBOR rate). For simplicity we consider a nominal amount $N = 1$ and regular quarterly payments such that $\delta_i = 1/4$ for all i.

Suppose the contract is agreed at time $t = 0$. The value of the contract at time $0 \le t < T_0$ to the party receiving the fixed payments is given by the formula

$$V(t) = -p(t, T_0) + p(t, T_n) + \frac{K}{4} \sum_{i=1}^{n} p(t, T_i),$$

where $p(t, T)$ denotes the price of a zero-coupon bond with maturity T at time t; see Brigo and Mercurio (2006, Section 1.5).

a) How is K determined if the contract should be fair when it is agreed upon?

b) Derive the mapping for the interest-rate swap in the discrete-time form $V_t = g(\tau_t, \boldsymbol{Z}_t)$ where $\boldsymbol{Z}_t$ are risk factors which you should identify, $\tau_t = t(\Delta t)$ and Δt is the length of the risk management time horizon.

c) Hence find the linearized loss operator for the interest-rate swap.

Exercise 9.9 (Mapping a currency forward)

Denote by (e_t) the exchange rate between two countries (foreign and domestic), that is at time T one unit of the foreign currency can be exchanged for e_T units of the domestic currency. As a concrete example, assume that e_t is the USD/EUR exchange rate, so that the foreign bond is a US treasury bond and the domestic bond a Euro area bond. A *currency* or *FX forward* is an agreement between two parties to buy/sell a prespecified amount N of USD at a future time point $T > t$ for a prespecified exchange rate $\bar{e}$ (and not at the market rate e_T). The future buyer is said to hold a long position, the other party is said to hold a short position in the contract.

a) Explain why the value of the contract at maturity T is equal to $V_T = N(e_T - \bar{e})$ (from the viewpoint of the party holding a long position), and show that a long position in the forward is equal to a portfolio of N US treasury bonds $p^{\mathrm{f}}(t, T)$ and $-N\bar{e}$ units of the Euro area zero coupon bond $p^{\mathrm{d}}(t, T)$.

b) Derive the mapping for the currency forward using as risk factors the yield to maturity of the Euro area and of the US treasury bond and the logarithmic exchange rate.

c) Compute the linearized loss operator for the mapping derived in b).

Exercise 9.10 (Asset liability management using duration and convexity)
Consider a pension fund which makes payments of 1M at the end of each of the next five years, that is for $T \in \mathcal{T} = \{1, 2, 3, 4, 5\}$; the current time is $t = 0$. Assume that the yield curve is flat with yield to maturity $y(0, T) = 1\%$ for $T \in \mathcal{T}$. Suppose that the capital of the fund today is exactly equal to the present value of the liabilities. The fund manager compares the following two strategies:

1) Perfect matching: Here the manager buys for each maturity $T \in \mathcal{T}$ zero-coupon bonds with nominal value 1M.
2) Immunization with maximal convexity: Here the manager invests only in zero bonds with maturity 1 and 5. The nominal values λ_1 and λ_5 are determined from the condition that in $t = 0$ the value of the assets should be equal to the value of the liabilities and that the duration of the cash flows on the asset side should be equal to the duration of the cash flows on the liability side.

a) Use numerical root finding to show that $\lambda_1 \approx 2.48$M and $\lambda_5 \approx 2.53$M for Strategy 2.
b) Compute the convexity of both strategies.
c) Compute for both strategies the change in the overall value of the position (assets $-$ liabilities) if i) all yields rise to 2% and ii) all yields drop to 0.1%.
d) In view of your results in b) and c), comment on the claim that Strategy 2 (which has higher convexity) outperforms Strategy 1.

Exercise 9.11 (Variance-covariance method for multivariate *t* risk-factor changes)
Suppose that the vector of risk factor changes satisfies $\boldsymbol{X} = (X_1, \dots, X_d) \sim t_d(\nu, \mu, \Sigma)$.

a) Explain why this might be a better model for market risk than a multivariate normal distribution.
b) Suppose that the linearized loss $L^{\Delta} = \lambda_0 + \boldsymbol{\lambda}'\boldsymbol{X}$ is a sufficiently accurate approximation of the true loss. Use Exercises 6.20 a) and 2.16 to derive formulas for $\mathrm{VaR}_\alpha(\lambda_0 + \boldsymbol{\lambda}'\boldsymbol{X})$ and, for $\nu > 1$, $\mathrm{ES}_\alpha(\lambda_0 + \boldsymbol{\lambda}'\boldsymbol{X})$.
c) Use b) to calculate $\mathrm{VaR}_{0.99}(L^{\Delta}_{t+1})$ and $\mathrm{ES}_{0.99}(L^{\Delta}_{t+1})$ for the linearized portfolio loss over a one-day horizon for a portfolio of two stocks with portfolio weights $w_1 = 0.7$ and $w_2 = 0.3$ and a value today of $V_t = 1$ (in M EUR). You should assume that the log-returns are bivariate t distributed with annualized volatilities $\sigma_1 = 0.2$, $\sigma_2 = 0.25$, correlation $\rho = 0.4$ and degrees of freedom $\nu = 5$. Assume that the expected returns are 0 and that one year consists of 250 trading days. Use the square-root-of-time scaling to compute daily volatilities.
d) Compare your results with Exercise 2.17 and verify that they are consistent if you allow the degree of freedom to tend to infinity.

Exercise 9.12 (Sampling from an empirical distribution function)
Let $L_1, \dots, L_n$ denote realizations of $L \sim F$ and let

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{L_i \le x\}}, \quad x \in \mathbb{R},$$

be the empirical distribution function based on $L_1, \dots, L_n$. Show that sampling from F_n corresponds to randomly drawing from $L_1, \dots, L_n$ with replacement.

Exercise 9.13 (Square-root-of-time rule for serially independent risks)

a) Assume the losses for the next h days can be modelled as independent $L_t \sim \mathrm{N}(\mu, \sigma^2)$, $t \in \{1, \dots, h\}$. Show that the total loss $L_1 + \dots + L_h$ satisfies the rule

$$\mathrm{VaR}_\alpha(L_1 + \dots + L_h) = (h - \sqrt{h})\mu + \sqrt{h}\,\mathrm{VaR}_\alpha(L_1)$$

which, for $\mu = 0$, is known as the square-root-of-time rule.

Note. In MFE (2015, Example 9.4) a multivariate version of this result based on independent multivariate normal random vectors is given.

b) Suppose you hold 100 shares of a stock whose daily returns are independent normal distributed with mean 0.01 and standard deviation 0.2. What is VaR_α for a 9-day time horizon?

c) Consider independent losses $L_1, \dots, L_h$ from a distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 < \infty$. Show that for h large,

$$\mathrm{VaR}_\alpha(L_1 + \dots + L_h) \approx h\mu + \sqrt{h}\sigma\Phi^{-1}(\alpha).$$

Exercise 9.14 (Square-root-of-time rule for serially dependent, jointly normal risks)
Assume the losses for the next h days can be jointly modeled as $(L_1, \dots, L_h) \sim \mathrm{N}_h(\boldsymbol{\mu}, \Sigma)$, where $\boldsymbol{\mu} = \mu\mathbf{1} \in \mathbb{R}^d$ for $\mu \in \mathbb{R}$ and $\Sigma = \sigma^2 P$ for a correlation matrix P with equal off-diagonal entries $\rho \in [0, 1]$. Show that the total loss $L_1 + \dots + L_h$ satisfies the rule

$$\mathrm{VaR}_\alpha(L_1 + \dots + L_h) = \Big(h - \sqrt{h(1 + (h-1)\rho)}\Big)\mu + \sqrt{h(1 + (h-1)\rho)}\,\mathrm{VaR}_\alpha(L_1).$$

Note. For $\rho = 0$ (independent $L_1, \dots, L_h$), one obtains the (square-root-of-time) rule as in Exercise 9.13. For $\rho = 1$, we obtain $\mathrm{VaR}_\alpha(L_1 + \dots + L_h) = h\,\mathrm{VaR}_\alpha(L_1)$ which is clear since VaR_α is comonotone additive.

Exercise 9.15 (Tests based on VaR violations)
Suppose you are a regulator trying to assess the quality of the value-at-risk system used by a bank. The bank has estimated the one-day $\mathrm{VaR}_{0.99}$ of its trading book every day for two years (500 observations). The system has produced 14 $\mathrm{VaR}_{0.99}$ violations on days 7, 61, 70, 75, 90, 130, 200, 245, 367, 371, 385, 403, 406 and 487.

a) How many $\mathrm{VaR}_{0.99}$ violations would you expect in this period if the bank has a very high-quality forecast model for its trading losses?

b) Carry out a binomial test of the quality of the bank's forecast model and report the result.

c) Carry out a test of the hypothesis that the violations occur as an iid Bernoulli sequence.

Exercise 9.16 (Elicitability and comparison of VaR estimates)

The scoring function for $\mathrm{VaR}_\alpha(L)$ is given by $S^q_\alpha(y,l) = |I_{\{l\le y\}} - \alpha||l-y|$ where we recall that y stands for the forecast and l for the realized loss.

a) Use software to draw a graph of $S^q_\alpha(y,l)$ as a function of l for $y=1$ and $\alpha\in\{0.5,0.9\}$. Use the graph to explain intuitively why the minimizer of $y\mapsto \mathbb{E}(S^q_\alpha(y,L))$ for $\alpha=0.9$ is larger than for $\alpha=0.5$.

b) Table E.9.1 contains 12 realized one-period losses (generated from a GARCH(1, 1) model with standardized $t_{3.5}$ innovations) and corresponding $\mathrm{VaR}_{0.99}$ predictions $\widehat{\mathrm{VaR}}^t_{0.99,\mathrm{N}}$ and $\widehat{\mathrm{VaR}}^t_{0.99,t}$ based on $\mathrm{N}(0,1)$ innovations and fitted standardized t innovations, respectively. Use a suitable test statistic based on elicitability theory to compare the quality of the risk measure predictions and state whether the result is as you would expect.

t	1	2	3	4	5	6	7	8	9	10	11	12
L_{t+1}	−0.73	−2.35	1.30	0.05	−0.93	−0.81	−0.06	0.45	3.72	−0.47	1.21	−0.18
$\widehat{\mathrm{VaR}}^t_{0.99,\mathrm{N}}$	3.45	3.45	3.46	3.46	3.47	3.47	3.47	3.48	3.48	3.48	3.49	3.49
$\widehat{\mathrm{VaR}}^t_{0.99,t}$	3.76	3.73	3.70	3.67	3.64	3.61	3.58	3.56	3.53	3.50	3.47	3.45

Table E.9.1 Losses from a GARCH(1, 1) model with standardized $t_{3.5}$ innovations and corresponding $\mathrm{VaR}_{0.99}$ predictions $\widehat{\mathrm{VaR}}^t_{0.99,\mathrm{N}}$ and $\widehat{\mathrm{VaR}}^t_{0.99,t}$ based on $\mathrm{N}(0,1)$ innovations and fitted standardized t innovations, respectively.

Advanced

Exercise 9.17 (Delta and delta-gamma approximation)

The aim of this exercise is to verify the formulas for delta and delta-gamma approximations in MFE (2015, Section 9.1.2). The derivation of the first-order Taylor series approximation (delta approximation) and the second-order Taylor series approximation (delta-gamma approximation) of the one-period loss random variable

$$L_{t+1} = -(g(\tau_t+\Delta t, \boldsymbol{Z}_t+\boldsymbol{X}_{t+1}) - g(\tau_t,\boldsymbol{Z}_t))$$

is easier to carry out if one views the mapping $g:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}$ as a function of a single vector $\boldsymbol{y}=(\tau_t,\boldsymbol{z})$, so $g(\tau_t,\boldsymbol{z}) = g(\boldsymbol{y})$.

a) Derive the delta approximation L^Δ_{t+1} of L_{t+1} by using the first-order Taylor series approximation of $g(\boldsymbol{y}_1)$ about $\boldsymbol{y}_0$ for $\boldsymbol{y}_0=(\tau_t,\boldsymbol{z})$ and $\boldsymbol{y}_1=(\tau_t+\Delta t,\boldsymbol{z}+\boldsymbol{x})$.

b) Derive the delta-gamma approximation $L^{\Delta\Gamma}_{t+1}$ of L_{t+1} by using the second-order Taylor series approximation of $g(\boldsymbol{y}_1)$ about $\boldsymbol{y}_0$ for $\boldsymbol{y}_0 = (\tau_t, \boldsymbol{z})$ and $\boldsymbol{y}_1 = (\tau_t + \Delta t, \boldsymbol{z} + \boldsymbol{x})$.

Exercise 9.18 (Standard methods for market risk)

The aim of this exercise is to compare the various statistical methods for market risk management discussed in MFE (2015, Section 9.2) in the case of a simple portfolio of two stocks. Consider BMW and Siemens (SIE) (which were also analysed in MFE (2015, Section 3.2.1)) from 3 January 2000 to 31 December 2009; see the dataset `EURSTX_const` of the R package `qrmdata` for the data. You should work through the following steps:

a) Load the data and plot the stock prices for the period in question. Also compute and plot the log-returns and construct a scatterplot of one series against the other.

b) Consider a portfolio consisting of BMW and SIE stocks in the ratio 1:10. Write a function to evaluate the corresponding loss operator (in other words, a function which takes the matrix of risk-factor changes and the portfolio weights and returns the resulting losses).

c) Implement as many of the following methods as you can:
   i) the variance-covariance method;
   ii) the standard historical simulation method;
   iii) the Monte Carlo method assuming a multivariate t distribution for the log-returns (you will need a function to estimate a bivariate t distribution such as `fit.mst()` in the R package `QRM`);
   iv) the method which uses a GPD approximation to the tail of the distribution of historically simulated losses (you will need a function to fit a GPD model to excess losses over a threshold such as `fit_GPD_MLE()` in the R package `qrmtools`).

d) Compute $\mathrm{VaR}_{0.99}$ and $\mathrm{ES}_{0.99}$ estimates for each of the methods; for the Monte Carlo method, use a sample size of 10^4.

e) Plot a histogram of the losses and display the computed $\mathrm{VaR}_{0.99}$ and $\mathrm{ES}_{0.99}$ estimates (for example with vertical lines). Briefly compare the values of the estimates.

Exercise 9.19 (Standard methods for an option position)

Suppose a bank has a position consisting of a short put option on the S&P 500 index and a long position $\lambda > 0$ in the stock with total value in discrete-time mapping notation given by

$$V_t = \lambda S_t - P^{\mathrm{BS}}(\tau_t, S_t; r, \sigma_t, K, T),$$

where P^{BS} denotes the Black-Scholes put option formula with the usual terms (see MFE (2015, Example 9.1) for a similar situation with a call option). Suppose the option was sold on the 4 June 2009 and that this is day $t = 0$. On this day $\tau_0 = 0$, the SP500 value is $S_0 = 942.46$ and the value of the VIX volatility index is $\sigma_0 = 30.18$. Consider an option with maturity $T = 5$ years and strike $K = 1050$. Assume that there is a fixed interest rate of $r = 1\%$ per annum.

a) Determine λ if $V_0 = 0$; a function for calculating the value of a put option can be found in the R package `qrmtools`.

b) Taking the log stock price $\log S_t$ and the volatility σ_t as the risk factors derive the linear and quadratic loss operators expressing changes in the value of the position in terms of changes in the value of these risk factors.

c) Suppose that you have to measure the market risk of the position at the end of the day on 4 June 2010, one year after the option was sold. On this day we have $\tau_t = 1$, $S_t = 1064.88$ and $\sigma_t = 35.48$. The data can be found in the datasets `SP500` and `VIX` of the R package `qrmdata`. Use the period from 14 June 2006 to 4 June 2010 (1001 values). Estimate the 99% one-day value-at-risk and expected shortfall for the position using the following methods and summarize the differences in the results:

   i) the variance-covariance method;

   ii) the standard historical simulation method with full revaluation of the position;

   iii) the standard historical simulation method with a delta-gamma approximation for the change in the value of the position;

   iv) the univariate dynamic historical simulation method described in MFE (2015, Section 9.2.4); see the R package `rugarch`.

Exercise 9.20 (Elicitability of the exponential risk measure)
For $\alpha > 0$ and L with $\mathbb{E}(\exp(\alpha L)) < \infty$, the *exponential risk measure* $\varrho_\alpha^{\exp}(L)$ for the loss function $\ell(x) = \exp(\alpha x)$ is given by the solution of the equation (in m)

$$\mathbb{E}\big(1 - \exp(\alpha(L - m))\big) = 0, \tag{E1}$$

which leads to $\varrho_\alpha^{\exp}(L) = \frac{1}{\alpha}\log \mathbb{E}(\exp(\alpha L))$; see MFE (2015, Example 8.8) for details (we consider $c = 0$ for simplicity). Show that $\varrho_\alpha^{\exp}(L)$ is elicitable and give a consistent scoring function.

Hint. For $s(m, l) = 1 - \exp(\alpha(l - m))$, $m, l \in \mathbb{R}$, (E1) can be written as $\mathbb{E}(s(m, l)) = 0$. Use this to show that a consistent scoring function is given by $S(y, l) = \int_l^y s(m, l)\,\mathrm{d}m$.

Exercise 9.21 (Accuracy of the empirical quantile estimator)
In this exercise we will attempt to understand the magnitude of the error when an empirical quantile is used to estimate value-at-risk. Let $U_1, \dots, U_n \overset{\text{ind.}}{\sim} \mathrm{U}(0, 1)$ and let the corresponding order statistics be denoted $U_{(1)} \le U_{(2)} \le \dots \le U_{(n)}$. Let F_X be a continuous and strictly increasing distribution function and define $X_i = F_X^{\leftarrow}(U_i)$, $i \in \{1, \dots, n\}$, with corresponding order statistics $X_{(1)} \le \dots \le X_{(n)}$.

a) Let $N \sim \mathrm{B}(n, p)$ for $n \in \mathbb{N}$ and $p \in [0, 1]$. Let $G(x; a, b)$ denote the distribution function of a $\mathrm{Beta}(a, b)$ distribution with parameters $a > 0$ and $b > 0$. Show that the distribution function of N satisfies $F_N(x; n, p) = \bar{G}(p; \lfloor x \rfloor + 1, n - \lfloor x \rfloor)$.

   Note. This is how R evaluates `pbinom(x, size = n, prob = p)`.

b) Hence show that $U_{(k)}$ has a $\mathrm{Beta}(k, n - k + 1)$ distribution.

c) Derive an expression for $F_{X_{(k)}}^{\leftarrow}(\alpha)$ in terms of F_X and the quantiles of a beta distribution.

d) For notational convenience, let $q_X(u) = F_X^{\leftarrow}(u)$ and $\theta_{(k)} = \mathbb{E}U_{(k)}$, $k \in \{1, \dots, n\}$. According to Taylor's Theorem, for $u \in (0,1)$ close to $\theta_{(k)}$, it is known that

$$q_X(u) \approx q_X(\theta_{(k)}) + q'_X(\theta_{(k)})(u - \theta_{(k)}). \tag{E1}$$

Find approximations for $\mathbb{E}(X_{(k)})$ and $\mathrm{var}(X_{(k)})$ in terms of $q_X(\theta_{(k)})$, $q'_X(\theta_{(k)})$, k and n.

Exercise 9.22 (PCA factor model for US zero-coupon bond yields)
Repeat the analysis of MFE (2015, Example 9.3) for daily US treasury zero-coupon bond yield data. Work through the following steps:

a) Consider daily US treasury ZCB yields with maturities from one to 30 years from 2 January 2002 to 30 December 2011.
b) Apply principal component analysis (PCA) to the daily yield changes and determine how much of the total variability can be explained by the first three sample principal components.
c) Plot each of the loadings corresponding to the three largest eigenvalues as a function of time and compare with MFE (2015, Figure 9.3).
d) Plot the time series of the first three sample principal components.

The data can be found in the dataset `ZCB_USD` of the R package `qrmdata`.

10 Credit Risk

Review

Exercise 10.1 (Management of counterparty risk)
Explain the notion of counterparty risk and the challenges arising in the management of this risk type. Describe potential strategies for managing counterparty risk.

Exercise 10.2 (Bond investment under decreasing credit quality)
Consider a bond investor who expects the credit quality of bond issuer A to decrease. Develop a strategy to profit from this view by using CDS contracts with reference entity A and discuss the risks of the strategy.

Exercise 10.3 (Rating transition matrices)
A bank uses a simple internal rating system in which there are only two ratings, A and B, as well as a default state D. Suppose the one-year transition probabilities are given as follows, where the (i,j)th entry denotes the probability of migrating from rating state i to j, with a few missing entries:

	A	B	D
A	0.80	0.15	?
B	0.10	?	0.20
D	?	?	?

Complete the table of transition probabilities and calculate the probabilities that A-rated and B-rated obligors default over a two-year period.

Exercise 10.4 (Credit risk in the Merton model)

a) Explain the modelling of the default of a firm in the Merton model and discuss the relationship between equity and debt and European options on the asset value of the firm.
b) Consider a firm whose asset value follows a geometric Brownian motion with drift $\mu_V = 0.1$ and volatility $\sigma_V = 0.2$. Assume that the current ($t = 0$) value of the assets is equal to 200, that the nominal value of the liabilities is equal to 100 (all in M EUR) and that the maturity of the liabilities is in $T = 1$ (year). Calculate the default probability of the firm.
c) Discuss strengths and weaknesses of the Merton model as a typical structural model. Consider in particular the properties of credit spreads.
d) How does the price of the risky debt react to an increase in the volatility σ_V of the asset value of the firm? How is this related to the incentives of bondholders?

Exercise 10.5 (Market-based credit risk management and procyclicality)
Why might the use of a credit risk management system whose main input are market prices such as the public-firm EDF model described in MFE (2015, Section 10.3.3) lead to procyclical capital requirements and what are the resulting economic problems?

Exercise 10.6 (Limitations of risk-neutral pricing)
Explain the pros and cons of risk-neutral pricing for credit products and contrast risk-neutral pricing with the more traditional pricing methodology for loans. Which approaches would you use for the following products?

a) A retail loan to a small or medium sized company.
b) An option that gives the holder the right to enter into a CDS contract on a major corporation at a future date for a spread fixed today.

Exercise 10.7 (Models with stochastic hazard rate)
Explain why a simple hazard-rate model is not suitable for the computation of VaR for corporate bonds.

Basic

Exercise 10.8 (Risk-neutral valuation)
Consider a simple one-period model for the price of a defaultable zero-coupon bond with nominal value 1, maturity $T = 1$ (year) and with deterministic recovery rate $1 - \delta = 0.4$. The default probability of the bond is $1 - p = 0.01$, the risk-free simple interest rate is 0.025, and the current ($t = 0$) price of the bond is $p_1(0, 1) = 0.961$. The price of the bond in $T = 1$ is thus either 1 (in the case of no default) or 0.4 (in the case of default).

a) Compute the expected payoff of the bond.
b) Determine the corresponding risk-neutral default probability.
c) Compute the price and replicating portfolio for a stylized credit default swap with payoff $I_{\{\tau \le T\}}$, where τ denotes the time of default of the bond.

Exercise 10.9 (Estimating transition probabilities for a discrete-time Markov chain)
Let $S = \{0, \dots, n\}$ be the set of rating states of increasing creditworthiness with 0 representing default. For $t = 0, \dots, T - 1$, $j \in S \setminus \{0\}$ and $k \in S$, let N_{tj} denote the number of companies that are rated j at time t and which are still in the rating system at time $t + 1$, and let N_{tjk} denote the number of those companies that are rated k at time $t+1$. A discrete-time, stationary Markov chain is fitted to the data (N_{tj}) and (N_{tjk}).

Show that the maximum likelihood estimator of the transition probability p_{jk} is given by

$$\hat{p}_{jk} = \frac{\sum_{t=0}^{T-1} N_{tjk}}{\sum_{t=0}^{T-1} N_{tj}}.$$

Hint. Use the fact that conditional on N_{tj}, the numbers migrating to each state k have a multinomial distribution.

Exercise 10.10 (Default probability in Merton's model)
In Merton's model assume that the growth rate of assets μ_V is positive and that the initial asset value V_0 exceeds the liability B. Show that the probability of default at the horizon date T is an increasing function of the volatility σ_V.

Exercise 10.11 (The Pareto survival model)
The Pareto distribution with shape parameter α and scale parameter κ, denoted $\mathrm{Pa}(\alpha, \kappa)$, has survival function

$$\bar{F}(t) = \left(\frac{\kappa}{\kappa + t}\right)^{\alpha}, \quad t \geq 0,\ \kappa > 0,\ \alpha > 0;$$

provided that $\alpha > n$, the moments of $\tau \sim \mathrm{Pa}(\alpha, \kappa)$ are given by $\mathbb{E}(\tau^n) = \kappa^n n! / \prod_{i=1}^{n}(\alpha - i)$.

a) Derive the hazard function $\gamma(t)$ of this distribution. Is the hazard rate increasing or decreasing in t?
b) Show that $\tau \sim \mathrm{Pa}(\alpha, \kappa)$ can be considered as a mixture model with mixing variable $\Lambda \sim \mathrm{Ga}(\alpha, \kappa)$ where, given $\Lambda = \lambda$, τ has an exponential distribution with parameter λ.
c) Suppose $\alpha > 2$ and consider an exponential random variable $\tilde{\tau}$ with parameter $\lambda = (\alpha - 1)/\kappa$. Show that $\mathbb{E}(\tau) = \mathbb{E}(\tilde{\tau})$ but $\mathrm{var}(\tau) > \mathrm{var}(\tilde{\tau})$.

   Note. This illustrates that mixing over λ leads to greater variation in the realizations of τ.

Exercise 10.12 (The Gompertz survival model)
The Gompertz model is widely used by actuaries in mortality modelling. The distribution function is given by

$$F(t) = 1 - \exp\Big(-\frac{a}{b}\Big(e^{bt} - 1\Big)\Big), \quad t \geq 0,\ a > 0,\ b > 0.$$

a) Caculate the hazard function and the cumulative hazard function of this distribution function.
b) The Gompertz–Makeham model is an extension of the Gompertz model. If the hazard function of the Gompertz distribution is $\gamma_{\mathrm{G}}(t)$, the Gompertz–Makeham distribution has hazard function $\gamma_{\mathrm{GM}}(t) = \gamma_{\mathrm{G}}(t) + c$ for some constant $c > 0$. Calculate the distribution function of this model.

Exercise 10.13 (Reduced-form credit spread under different recovery assumptions)
Suppose that under the risk neutral measure $\mathbb{Q}$, the default time τ follows a hazard rate model with continuous hazard function $\gamma^{\mathbb{Q}}$. Consider a defaultable zero coupon bond with maturity date T and a valuation date $t < T$ such that $\tau > t$. Derive a formula for the price $p_1(t, T)$ and the credit spread $c(t, T)$ of a defaultable zero coupon bond in the following recovery models:

a) RT (recovery of treasury);

b) RF (recovery of face value).

In both cases, give the limit of the spread for the case where T converges to t.

Exercise 10.14 (Pricing and calibration of a CDS)
Consider a CDS on firm R with time to maturity $T = 5$ years and quarterly premium payments. Assume that the risk free continuously compounded interest rate is equal to $r > 0$ and that under the risk-neutral measure $\mathbb{Q}$ the default time τ_R of the reference entity follows a hazard-rate model where the hazard rate is constant and equal to $\bar{\gamma}^{\mathbb{Q}} > 0$. Assume moreover that the loss-given-default of the reference entity is given by the constant $\delta \in (0, 1]$.

a) Compute the value in $t = 0$ of the CDS cash-flows from the viewpoint of the protection seller for a generic annualized CDS spread x.

b) Suppose now that the annualized spread of the CDS observed on the market is $x^* = 42\,\text{bp}$, that the loss-given-default is $\delta = 0.6$ and that $r = 0.02$. Calibrate a constant hazard model for τ_R under the risk-neutral probability measure $\mathbb{Q}$. Use both the approximate formula $\bar{\gamma}^{\mathbb{Q}} \approx x^*/\delta$ and exact calibration (numerical root finding is required for the latter).

c) Can the value of $\bar{\gamma}^{\mathbb{Q}}$ computed as in b) be used for computing the VaR of a credit exposure towards R such as a loan?

Exercise 10.15 (Application of the Feynman–Kac formula)
Use the Feynman–Kac formula, see MFE (2015, Lemma 10.21), to solve the following two terminal value problems. In both cases $\mu \in \mathbb{R}$ and $\sigma > 0$ are given constants.

a) $f_t(t,\psi) + \mu f_\psi(t,\psi) + \frac{1}{2}\sigma^2 f_{\psi\psi}(t,\psi) = 0$, $(t,\psi) \in [0,T) \times \mathbb{R}$, with terminal condition $f(T,\psi) = \psi^2$.

b) $f_t(t,\psi) + \mu\psi f_\psi(t,\psi) + \frac{1}{2}\sigma^2\psi^2 f_{\psi\psi}(t,\psi) = rf(t,\psi)$, $(t,\psi) \in [0,T) \times \mathbb{R}_+$ with terminal condition $f(T,\psi) = \log\psi$.

Exercise 10.16 (Vasicek model)
Consider the stochastic differential equation (SDE) for an Ornstein–Uhlenbeck process,

$$\mathrm{d}\psi_t = a(b - \psi_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \quad a, b > 0.$$

This SDE is sometimes used as a model for the short-rate of interest known as the *Vasicek model*. Define the function $F(t,\psi) = \mathbb{E}(\exp(-\int_t^T \psi_s\,\mathrm{d}s) \mid \psi_t = \psi)$, the price of a default-free zero coupon bond. Conjecture that $F(t,\psi) = \exp(\alpha(t,T) + \beta(t,T)\psi)$ and derive a system of ordinary differential equations for α and β.

Exercise 10.17 (Simplified CDS pricing)
In this exercise we consider an approximation that reduces CDS pricing to the pricing of survival claims. Consider a CDS with deterministic loss-given-default δ and premium payment dates $t_1 < t_2 < \cdots < t_N = T$, and let $t_0 = 0$. We slightly modify the standard definition of the payoff of the default payment leg and assume that if $t_i < \tau_R \le t_{i+1}$ for any $0 \le i \le N-1$, the protection buyer receives the amount δ at time t_{i+1} (and not directly at τ_R).

a) Denote by $\mathbb{Q}$ the risk-neutral measure and by (r_t) the short rate of interest. Show that the price at (for simplicity) $t = 0$ of the modified default payment leg equals

$$\delta \sum_{i=1}^{N} \mathbb{E}^{\mathbb{Q}}\Big(e^{-\int_0^{t_i} r_s \,\mathrm{d}s}(I_{\{\tau_R > t_{i-1}\}} - I_{\{\tau_R > t_i\}})\Big). \tag{E1}$$

Suppose that (r_t) is a positive process. Is (E1) bigger or smaller than the value of the standard default payment where the default payment is made directly at τ_R?

b) Suppose that the short rate is deterministic and let $p_0(t,T) = e^{-\int_0^t r(s)\,\mathrm{d}s}$. Show that in this case, (E1) can be written as

$$\delta\Big(p_0(0,t_1) - p(0,t_N)\mathbb{Q}(\tau_R > t_N) + \sum_{i=1}^{N-1} \mathbb{Q}(\tau_R > t_i)(p_0(0,t_{i+1}) - p_0(0,t_i))\Big). \tag{E2}$$

c) Use the approximation formula (E2) to compute numerically the fair spread of a CDS with maturity $T = 5$ years and quarterly premium payments. Assume that under $\mathbb{Q}$ the default intensity follows a Cox–Ingersoll–Ross (CIR) process with dynamics $\mathrm{d}\Psi_t = \kappa(\theta - \Psi_t)\,\mathrm{d}t + \sigma\sqrt{\Psi_t}\,\mathrm{d}W_t$ with parameters $\Psi_0 = 0.03$, $\kappa = 1$, $\theta = 0.04$ and $\sigma = 0.2$, and assume $r(t) \equiv 0.01$, $\delta = 0.6$. How does an increase in Ψ_0, θ or σ affect the spread?

Advanced

Exercise 10.18 (Matrix exponential for continuous-time Markov chain)

The matrix exponential of the Markov chain generator $\Lambda \in \mathbb{R}^{(n+1)\times(n+1)}$ for some $n \in \mathbb{N}$ can be calculated using a number of software packages. One case where it can be calculated by simple matrix multiplication occurs when the generator is diagonalizable, meaning that there exists an invertible matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$ such that $A^{-1}\Lambda A = D$ where $D = \operatorname{diag}(d_0, \ldots, d_n)$ is a diagonal matrix containing the eigenvalues of Λ. Show that, in this case, the matrix of transition probabilities $\boldsymbol{P}(t)$ for the interval $[0,t]$ may be written as

$$\boldsymbol{P}(t) = A \operatorname{diag}(e^{d_0 t}, \ldots, e^{d_n t})A^{-1}.$$

Exercise 10.19 (Calibrating hazard rate model when CDS spread curve is flat)

Suppose that the positive constant $\bar{\gamma}^{\mathbb{Q}}$ satisfies

$$x^* \Delta t e^{-(r+\bar{\gamma}^{\mathbb{Q}})\Delta t} = \delta \bar{\gamma}^{\mathbb{Q}} \int_0^{\Delta t} e^{-(r+\bar{\gamma}^{\mathbb{Q}})t}\,\mathrm{d}t. \tag{E1}$$

a) Show that $\bar{\gamma}^{\mathbb{Q}}$ must also satisfy

$$x^* \Delta t \sum_{k=1}^{N} e^{-(r+\bar{\gamma}^{\mathbb{Q}})k\Delta t} = \delta \bar{\gamma}^{\mathbb{Q}} \int_0^{N\Delta t} e^{-(r+\bar{\gamma}^{\mathbb{Q}})t}\,\mathrm{d}t \quad \text{for all} \quad N \in \mathbb{N}.$$

b) Conclude that, when the CDS spread curve is flat, when interest rates are constant, and when premium payments follow a regular schedule, a constant risk-neutral hazard rate $\bar{\gamma}^{\mathbb{Q}}$ can be determined from a quoted market spread x^* by solving (E1).

Exercise 10.20 (Risk management for a corporate bond)
Consider a single corporate zero-coupon bond with time to maturity $\overline{T} = 2$ years and denote by τ the default time of the bond. The recovery value of the bond is zero so that its payment at $\overline{T}$ is $I_{\{\tau>\overline{T}\}}$ and the current price of the bond is p_0. Assume that the risk-free interest rate is constant and equal to $r > 0$, that under the historical probability measure P, τ is doubly stochastic with hazard rate process $\gamma^{\mathbb{P}}$ and that under the risk-neutral measure $\mathbb{Q}$, τ is doubly stochastic with hazard rate process $\gamma^{\mathbb{Q}}$. Moreover,

$$\gamma_t^{\mathbb{P}} = \psi_t \text{ and } \gamma_t^{\mathbb{Q}} = 2\psi_t,$$

where (ψ_t) follows a CIR process with $\mathbb{P}$-parameters $\kappa^{\mathbb{P}}, \theta^{\mathbb{P}}, \sigma > 0$ and $\mathbb{Q}$-parameters $\kappa^{\mathbb{Q}}, \theta^{\mathbb{Q}}, \sigma > 0$.

a) Develop a simulation algorithm to compute the loss distribution and the VaR of the bond over the time horizon $T = 1$ year. Define for this the function

$$p(t_1, t_2, \psi; r, \rho, \kappa, \theta, \sigma) = \mathbb{E}\left(\exp\left(-\int_{t_1}^{t_2} r + \rho\psi_s \,\mathrm{d}s\right) \,\middle|\, \psi_{t_1} = \psi\right),$$

where (ψ_t) follows a CIR process with generic parameters κ, θ, σ. Note that the function $p(t_1, t_2, \psi; r, \rho, \kappa, \theta, \sigma)$ is known explicitly, see MFE (2015, Section 10.6.2), but its precise form is not required in this exercise. Explain, for which part of the simulation algorithm risk-neutral, respectively historical, quantities are needed.

b) Modify the algorithm proposed in a) so that it can be used for pricing a put option on the bond with maturity $T = 1$ and exercise price $K = p_0$ via Monte Carlo simulation.

Exercise 10.21 (Credit value adjustment formula)
The theoretical background for this exercise is discussed in MFE (2015, Section 17.2), but it can be solved with a basic understanding of counterparty risk.

Consider two parties S and B who enter into a contract where the protection seller S provides protection against an adverse event and the protection buyer B pays premia in return. A case in point would be a credit default swap between protection seller S and protection buyer B or a reinsurance treaty between the primary insurer B and the reinsurer S. The market value of this contract at time t from the viewpoint of B is denoted V_t and the maturity of the contract is T. To account for the possibility that S might default before T, a value adjustment is computed. Under the simplifying assumption that the default time of S and the counterparty-risk free price process (V_t) are independent and this value adjustment is given by

$$\mathrm{CVA}^{\mathrm{indep}} = \delta^{\mathrm{S}} \int_0^T e^{-rt} \mathbb{E}^{\mathbb{Q}}(V_t^+) f_{\mathrm{S}}(t) \,\mathrm{d}t.$$

Here δ^{S} gives the loss given default of S; $r \geq 0$ is the risk-free interest rate, and f_{S} denotes the density of the default time τ_{S} of S.

a) Explain the formula for $\mathrm{CVA}^{\mathrm{indep}}$.

b) Discuss the assumption that the market value (V_t) of the contract and the default time τ_{S} are independent for the special case of a reinsurance treaty between primary insurer B and reinsurer S.

c) Evaluate the formula for $\mathrm{CVA}^{\mathrm{indep}}$ in the case where V_t is normally distributed under $\mathbb{Q}$ with mean 0 and variance $\sigma^2 t^2$. Assume that, under $\mathbb{Q}$, τ_{S} is exponentially distributed with parameter γ_{S}.

   Hint. For $\alpha > 0$, the indefinite integral of $xe^{-x^2/(2\alpha^2)}$ is given by $-\alpha^2 e^{-x^2/(2\alpha^2)}$. For $\gamma > 0$, the integral of $te^{-\gamma t}$ is given by $-(1/\gamma)e^{-\gamma t}(t + 1/\gamma)$.

11 Portfolio Credit Risk Management

Review

Exercise 11.1 (Dependence in credit risk management)

a) Give economic reasons why it is generally not appropriate to assume independence between the default of different obligors in a loan portfolio.

b) How does dependence between defaults affect the loss distribution in a typical loan portfolio? Discuss implications for managing the risk of loan portfolios.

Exercise 11.2 (Gaussian threshold models)

Consider a Gaussian threshold model $(\boldsymbol{X}, \boldsymbol{d})$ where the critical variables $\boldsymbol{X}$ follow a factor model.

a) Describe the mathematical structure of such a model and explain the advantages of using a factor model for $\boldsymbol{X}$.

b) Explain the difference between default correlation and asset correlation in this context .

c) How is the threshold d_i chosen in order to calibrate the model to a given default probability p_i for obligor i?

d) Derive the form of an equivalent Bernoulli mixture model.

Exercise 11.3 (Tail dependence in threshold models)

Consider a threshold model $(\boldsymbol{X}, \boldsymbol{d})$ and explain why lower tail dependence of the critical variables might have a substantial impact on the tail of the credit loss distribution.

Exercise 11.4 (Exchangeable Bernoulli mixture models)

Consider an exchangeable Bernoulli mixture model for m obligors and denote by the random variable Q the conditional default probability.

a) Derive the relationship between the unconditional default probability π and the default correlation ρ_Y and the first two moments of Q. Why is ρ_Y always nonnegative in Bernoulli mixture models?

b) Explain why for large m the tail of the distribution of the number of defaults is ‘up to first order’ determined by the tail of Q.

Exercise 11.5 (Importance sampling in credit risk models)

a) Describe the basic idea of importance sampling and the motivation for using the method in credit risk.

b) Assume that a Bernoulli mixture with mixing variable $\boldsymbol{\Psi}$ is used to model the default state of a given credit portfolio. Why might importance sampling for the conditional loss distribution as described in MFE (2015, Algorithm 11.24) not suffice to obtain a low variance of the IS estimator, in particular for large portfolios.

Basic

Exercise 11.6 (Calculating moments of portfolio credit loss distributions)
Consider a portfolio containing $m = 1000$ equally rated credit risks. Assume that for every obligor $i \in \{1, \dots, m\}$ the exposure is $e_i = 1\text{M}\,\text{GBP}$ and the default probability is $p_i = 1\%$. Calculate the expected value and standard deviation of the portfolio loss in the following situations.

a) Loss-given-defaults (LGDs) are modeled as deterministic and $\delta_i = 0.4$, $i \in \{1, \dots, m\}$. Defaults are assumed to occur independently.

b) LGDs are modeled as deterministic as in a) but defaults are assumed to be dependent with pairwise default correlation (that is correlation between pairs of default indicators Y_i) given by $\rho = \operatorname{corr}(Y_i, Y_j) = 0.005$ for $i \neq j$.

c) Defaults are dependent as in b) but LGDs are modeled as random variables Δ_i satisfying $\Delta_i \sim \operatorname{Beta}(a, b)$, $i \in \{1, \dots, m\}$, where $a = 9.2$ and $b = 13.8$. Assume that LGDs are mutually independent across the portfolio and independent of the default indicator variables.

Hint. $\mathbb{E}(\Delta_1^k) = \prod_{l=0}^{k-1} \frac{a+l}{a+b+l}$, $k \in \mathbb{N}$.

Exercise 11.7 (Correlation bounds for default indicators)
Let $X_j \sim \operatorname{B}(1, p_j)$, $p_j \in (0, 1)$, $j \in \{1, 2\}$.

a) Show that the correlation coefficient ρ of (X_1, X_2) satisfies $\rho_{\min} \le \rho \le \rho_{\max}$ where

$$\rho_{\min} = \frac{W(p_1, p_2) - p_1 p_2}{\sqrt{p_1(1-p_1)p_2(1-p_2)}} \quad \text{and} \quad \rho_{\max} = \frac{M(p_1, p_2) - p_1 p_2}{\sqrt{p_1(1-p_1)p_2(1-p_2)}},$$

for the countermonotonicity copula W and the comonotonicity copula M.

Hint. Use Höffding's Theorem (MFE (2015, Theorem 7.28)).

b) Consider the case $p_1 = p_2 =: p$. For which value of p is the range of ρ maximal and what is this range?

Exercise 11.8 (Asset correlations in Gaussian threshold models with two-group structure)
Consider a Gaussian threshold model $(\boldsymbol{X}, \boldsymbol{d})$ in which the critical variables follow the factor model

$$X_i = \sqrt{\beta_i}\tilde{F}_i + \sqrt{1-\beta_i}\varepsilon_i, \quad i \in \{1, \dots, m\},$$

where $\tilde{F}_i, \varepsilon_1, \dots, \varepsilon_m$ are independent standard normal variables and $0 \le \beta_i \le 1$ for all i. The systematic variables $\tilde{F}_i$ in turn satisfy

$$\tilde{F}_i = \begin{cases} F_1, & i = 1, \dots, n, \\ \rho F_1 + \sqrt{1-\rho^2} F_2, & i = n+1, \dots, m, \end{cases}$$

where F_1, F_2 are independent standard normal factors, $0 \le \rho \le 1$ and $1 < n < m-1$. Obviously, this model has a two-group structure determining the dependence between defaults. Derive expressions for the within-group asset correlations (for each of the two groups) and the between-group asset correlation.

Exercise 11.9 (Joint default probabilities in Gaussian threshold models)
Consider a Gaussian threshold model $(\boldsymbol{X}, \boldsymbol{d})$ in which the critical variables follow the one-factor model

$$X_i = \sqrt{\beta_i} F + \sqrt{1-\beta_i} \varepsilon_i, \quad i \in \{1, \dots, m\}.$$

Here $F, \varepsilon_1, \dots, \varepsilon_m$ are independent standard normal variables and $0 \le \beta_i \le 1$ for all i. The goal of this exercise is to describe the equivalent Bernoulli mixture model.

a) For $1 \le k \le m$, express the joint probability that the group of obligors $\{i_1, \dots, i_k\} \subseteq \{1, \dots, m\}$ all default as a one-dimensional integral over the distribution of the factor F.
b) Show that in the exchangeable version of this model (that is for $d_i = d$ and $\beta_i = \rho = \operatorname{corr}(X_j, X_l)$ for all i, j, l) the formula in a) leads to

$$\pi_k = \int_{-\infty}^{\infty} \left(\Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{\rho} x}{\sqrt{1-\rho}} \right) \right)^k \phi(x) \, \mathrm{d}x;$$

that is the kth order joint default probability π_k is related to the default probability π of each obligor according to this formula.

Exercise 11.10 (Joint default probabilities in threshold models with Gumbel copula)
Derive a formula that relates higher-order default probabilities π_k and individual default probabilities π in an exchangeable default model of threshold type based on the Gumbel copula.

Exercise 11.11 (Exchangeable one-factor Bernoulli mixture models)
Suppose you have a portfolio of m similarly rated obligors. You decide to model the default state of this portfolio by an exchangeable Bernoulli mixture model with mixing variable Q. Recall that in this model, the default indicators $Y_j \in \{0, 1\}$ satisfy $\mathbb{P}(Y_j = y_j \,|\, Q = q) = q^{y_j}(1-q)^{1-y_j}$, $y_j \in \{0, 1\}$, $j \in \{1, \dots, m\}$, so that

$$\mathbb{P}(\boldsymbol{Y} = \boldsymbol{y} \,|\, Q = q) = \prod_{j=1}^{m} q^{y_j}(1-q)^{1-y_j} = q^{\sum_{j=1}^m y_j}(1-q)^{m - \sum_{j=1}^m y_j}.$$

a) Show that the kth order joint default probability $\pi_k = \mathbb{P}(Y_{j_1} = 1, \dots, Y_{j_k} = 1,\ 1 < j_1 < \dots < j_k < m)$ of any subgroup of $k \in \{1, \dots, m\}$ components defaulting satisfies $\pi_k = \mathbb{E}(Q^k)$, $k \in \{1, \dots, m\}$.

b) Compute the implied default correlation $\rho_Y = \mathrm{corr}(Y_{j_1}, Y_{j_2})$ for $j_1, j_2 \in \{1, \dots, m\}$, $j_1 \neq j_2$, in terms of π_1 and π_2.

Exercise 11.12 (Exchangeable beta mixture model)
Suppose the credit risk of a portfolio of m similar obligors is modelled with an exchangeable Bernoulli mixture model with mixing variable

$$Q \sim \mathrm{Beta}(a, b),$$

that is the mixing variable Q is distributed according to a beta distribution with parameters $a > 0$, $b > 0$ and density $f_Q(x) = x^{a-1}(1-x)^{b-1}/\beta(a, b)$, $x \in (0, 1)$, where $\beta(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ denotes the beta function.

a) Let $k \in \{1, \dots, m\}$. Compute the kth order joint default probability $\pi_k = \mathbb{P}(Y_{j_1} = 1, \dots, Y_{j_k} = 1)$ for all $1 < j_1 < \cdots < j_k \leq m$.

   Hint. $\Gamma(z + 1) = z\Gamma(z)$, $z > 0$.

b) Determine a, b such that the individual default probability π_1 is 1% and the default correlation ρ_Y is 0.5%.

c) Let $M = \sum_{j=1}^{m} Y_j$ denote the number of defaults in the portfolio. Show that its probability mass function $(p_k)_{k=0}^{m}$ is given by

$$p_k = \mathbb{P}(M = k) = \binom{m}{k} \frac{\beta(a + k, b + m - k)}{\beta(a, b)}, \quad k \in \{0, \dots, m\}.$$

d) Compute the probability mass function $\tilde{p}_k = \mathbb{P}(M = k)$, $k \in \{0, \dots, m\}$, under the assumption that the default indicators Y_j's are independent.

e) Suppose $m = 1000$. For $k \in \{0, \dots, 60\}$, plot p_k and $\tilde{p}_k$. Also plot $p_k/\tilde{p}_k$ and comment on the plot.

Exercise 11.13 (Exchangeable probit-normal mixture model)
Suppose the credit risk of a portfolio of m similar obligors is modelled with an exchangeable Bernoulli mixture model where the mixing variable Q has a probit-normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, that is

$$Q = \Phi(\mu + \sigma Z), \quad Z \sim \mathrm{N}(0, 1).$$

Therefore, $Q = \Phi(X)$ for $X \sim \mathrm{N}(\mu, \sigma^2)$ and thus, by Exercise 11.11 a), $\pi_k = \mathbb{E}(\Phi(X)^k)$, $k \in \{1, \dots, m\}$.

a) Derive the distribution function and density of Q.

b) Use Exercise 11.9 to infer that the exchangeable mixture model is equivalent to an exchangeable Gaussian threshold model with default probability π and asset correlation ρ which you should determine as functions of μ and σ.

c) Use these results to give the mean $\mathbb{E}(Q)$ of the probit-normal distribution.

Exercise 11.14 (Exchangeable Bernoulli mixtures with extremal mixing)
Consider an exchangeable Bernoulli mixture model for m obligors with given default probability $\mathbb{P}(Y_i = 1) = \pi \in (0, 1)$ but unspecified distribution function $G(q)$ of the mixing variable Q.

a) Determine the mixing distributions $G^{\min}(q)$ and $G^{\max}(q)$ such that the default correlation ρ_Y attains the bounds $\rho_Y^{\min} = 0$ respectively $\rho_Y^{\max} = 1$.
b) Show that for the mixing distribution $G^{\max}(q)$ the default indicators are comonotone.
c) Let $L = \sum_{i=1}^m Y_i$ (exposures are ignored for simplicity). Show that for the mixing distribution $G^{\max}(q)$ the expected shortfall $\mathrm{ES}_\alpha(L)$ is maximized over all mixing distributions $G(q)$ with $\int_0^1 q\,\mathrm{d}G(q) = \pi$ (that is all mixing distributions that are consistent with the constraint $\mathbb{P}(Y_i = 1) = \pi$). Does $G^{\max}(q)$ also maximize Value at Risk (proof or counterexample)?

Exercise 11.15 (Two-factor CreditRisk$^+$ model)
Consider a portfolio of $m = 1000$ obligors with an exposure of $e_i = 1\text{M}$ for each obligor $i \in \{1, \dots, m\}$. Suppose we model dependent defaults in the portfolio using a two-factor CreditRisk$^+$-style of model as in MFE (2015, Section 11.2.5). Assume that the default count variables $\tilde{Y}_i$, see MFE (2015, Section 11.2.5), satisfy

$$\tilde{Y}_i \,|\, ((\Psi_1, \Psi_2) = (\psi_1, \psi_2)) \sim \begin{cases} \mathrm{Poi}(0.01\psi_1), & i = 1, \dots, 500, \\ \mathrm{Poi}(0.01(0.5\psi_1 + 0.5\psi_2)), & i = 501, \dots, 1000. \end{cases}$$

Also assume that loss-given-default is 100% in all cases and that $\Psi_1, \Psi_2 \overset{\text{ind.}}{\sim} \mathrm{Ga}(1/2, 1/2)$; note that the second parameter $1/2$ is the rate parameter of the gamma distribution. Compute the mean and variance of the portfolio loss.

Exercise 11.16 (Large portfolio asymptotics application)
Over the years a retail banking division specialising in small commercial loans has had a consistent lending policy: 50% of its loans have been for the amount of 5M and 50% of its loans have been for the amount of 1M. Moreover, 50% of both the larger and smaller loans have been rated as ‘risky’ and have been assigned a default probability of 1% per annum, whereas the other 50% have been rated as ‘safe’ and have been assigned a default probability of 0.1% per annum.

The bank uses a one-factor Gaussian threshold model for its portfolio and carries out a fully internal calculation for economic capital purposes. In the one-factor model the risky loans are assumed to be 80% systematic (that is 80% of the variance of the driving ‘asset value’ variable is assumed to be explained by systematic factors) whereas the safe loans are assumed to be only 20% systematic. A deterministic loss-given-default of 0.6 is assumed.

The portfolio consists of $m = 10\,000$ individual loans and the bank decides to use a large portfolio argument to compute the 99.9% value-at-risk.

a) Derive the form of the asymptotic relative loss function $\bar{l}(\psi)$, see MFE (2015, Section 11.3.2), under the assumption that the portfolio is grown ad infinitum with the same lending policy.
b) Approximate the 99.9% VaR for the portfolio.

Exercise 11.17 (Large portfolio asymptotics with stochastic LGD)
Consider a Gaussian threshold model $(\boldsymbol{X}, \boldsymbol{d})$ with the critical variables following the one-factor model

$$X_i = \sqrt{\beta}_i F + \sqrt{1-\beta_i}\varepsilon_i, \quad i \in \{1, \dots, m\},$$

where $F, \varepsilon_1, \dots, \varepsilon_m \overset{\text{ind.}}{\sim} \mathrm{N}(0,1)$ and $0 \le \beta_i \le 1$ for all i. For $i = 1, \dots, m$, let $Y_i = I_{\{X_i \le d_i\}}$ denote the default indicator variables and $p_i = \mathbb{P}(Y_i = 1)$ the default probabilities.

a) Show that the model is equivalent to a one-factor Bernoulli mixture model for the default indicators where the common factor is $\Psi = -F$ and the conditional default probabilities take the form

$$p_i(\psi) = \mathbb{P}(Y_i = 1 \,|\, \Psi = \psi) = \Phi(\mu_i + \sigma_i \psi).$$

Also derive expressions for μ_i and σ_i.

b) Suppose there are $m = 10\,000$ obligors in the portfolio, half of them have exposure 2M GBP, default probability $p_i = 0.01$ and factor loading $b_i = 0.6$ and the other half of them have exposure 4M GBP, default probability $p_i = 0.05$ and factor loading $b_i = 0.8$. Assume the loss-given-default (LGD) is 0.6 for all obligors. Use a large portfolio argument to compute an approximation for the 99% value-at-risk of the portfolio loss.

c) Now suppose that a stochastic LGD Δ_i depending on the economic factor Ψ is introduced into the model for every obligor i. It is assumed that

   i) LGDs are conditionally independent given Ψ;

   ii) they are independent of the default indicators given Ψ; and

   iii) the expected LGD given Ψ satisfies $\mathbb{E}(\Delta_i \,|\, \Psi = \psi) = \Phi(0.5 + \psi)$.

   Recompute the approximate 99% value-at-risk to incorporate the stochastic LGDs.

Exercise 11.18 (Importance sampling with exponential tilting for gamma)
Let $X \sim \mathrm{Ga}(\alpha, \beta)$ be a gamma distributed random variable with corresponding density $f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x)$, $x > 0$, and moment-generating function $M_X(t)$.

a) Compute the exponentially tilted density $g_t(x) = \exp(tx) f_X(x)/M_X(t)$, $x > 0$. What are the constraints on the value of t?

b) Can the mean of X be shifted to arbitrary positive values under the importance sampling density?

Exercise 11.19 (Importance sampling with exponential tilting for Poisson)

a) Let $X \sim \mathrm{Poi}(\mu)$ under the probability measure $\mathbb{P}$. Determine the distribution of X under the probability measure $\mathbb{Q}$ obtained by exponential tilting; in other words, what is the distribution of X after exponential tilting?

b) Let $X_i \overset{\text{ind.}}{\sim} \mathrm{Poi}(\mu_i)$, $i = 1, \dots, m$, under the probability measure $\mathbb{P}$. We are interested in the distribution of the random variable $L = \sum_{i=1}^{m} e_i X_i$ where $e_1, \dots, e_m$ are known deterministic exposures. Suppose we change the probability measure to $\mathbb{Q}_t$ by exponentially tilting the random variable L. What is the joint distribution of $(X_1, \dots, X_m)$ under $\mathbb{Q}_t$?

c) Suppose we want to estimate the exceedance probability $\mathbb{P}(L > c)$ of L over some threshold $c \gg \mathbb{E}(L)$ with a Monte Carlo simulation. Since $c \gg \mathbb{E}(L)$, we would like to use as variance reduction technique importance sampling based on a proposal distribution $\mathbb{Q}_t$ under which L has mean c. Suppose this proposal distribution is obtained by exponentially tilting the distribution of L. What equation should we solve to determine the exponential tilting parameter t?

Advanced

Exercise 11.20 (Importance sampling for CreditRisk$^+$)

How are Exercises 11.18 and 11.19 relevant to the problem of using importance sampling to estimate tail probabilities of the loss distribution in CreditRisk$^+$?

Hint. You may assume the one-factor case.

Exercise 11.21 (Probability mass function of number of defaults under exchangeability)

Let $Y_1, \dots, Y_m \sim \mathrm{B}(1, p)$ be exchangeable default indicators of a portfolio of size m and let $M = \sum_{i=1}^{m} Y_i$ be the number of defaults in this portfolio. For $k \in \{1, \dots, m\}$, derive an expression for $\mathbb{P}(M = k)$ (that is for the probability that exactly k obligors default) in terms of the higher-order default probabilities $\pi_k, \dots, \pi_m$.

Note. Recall that π_k is the probability that an arbitrarily selected subgroup of at least k obligors default.

Exercise 11.22 (Threshold model with Archimedean survival copula)

Let $(\boldsymbol{X}, \boldsymbol{d})$ be an m-dimensional threshold model where $\boldsymbol{X}$ has an Archimedean survival copula with generator given by the Laplace transform $\widehat{G}$ of some distribution function G on $[0, \infty)$ with $G(0) = 0$.

a) Show that this model is equivalent to a one-factor Bernoulli mixture model with a factor Ψ whose distribution function has Laplace transform $\widehat{G}$ and show that the conditional default probabilities are given by

$$p_i(\psi) = 1 - \exp(-\psi \widehat{G}^{-1}(1 - p_i)), \quad i \in \{1, \dots, m\},$$

where p_i is the unconditional default probability.

b) Assume now that $\Psi \sim \mathrm{Ga}(\alpha, \alpha)$, $\alpha > 0$, so that $\widehat{G}$ is the Laplace transform of a gamma distribution. In this case, $\boldsymbol{X}$ has a so-called Clayton survival copula. Show that

$$p_i(\psi) = 1 - \exp\bigl(-\psi\alpha((1-p_i)^{-1/\alpha} - 1)\bigr), \quad i \in \{1, \dots, m\}.$$

c) Let $k_i = \alpha((1-p_i)^{-1/\alpha} - 1)$, $i \in \{1, \dots, m\}$. Show that $k_i \approx p_i$ for small p_i.

d) Explain why this threshold model is equivalent to the Bernoulli mixture model implied by a one-factor version of CreditRisk$^+$ with heterogeneous default rates.

e) Compute the joint default probability of obligors $i \neq j$ with identical unconditional default probabilities $p := p_i = p_j \in (0, 1)$. Work out how this probability behaves as $\alpha \to 0+$ and $\alpha \to \infty$. Which direction increases the amount of dependence in the model?

Exercise 11.23 (Copula of exchangeable CreditRisk$^+$ model)

Suppose that $\tilde{\boldsymbol{Y}} = (\tilde{Y}_1, \dots, \tilde{Y}_m)$ follows an exchangeable one-factor CreditRisk$^+$ model in which the $\tilde{Y}_i$ are conditionally independent and Poisson-distributed with parameter $c\Psi$ where $\Psi \sim \mathrm{Ga}(\alpha, \alpha)$ and $c > 0$, $\alpha > 0$. Let $Y_i = I_{\{\tilde{Y}_i > 0\}}$ be the default indicator for firm i. Show that for any $1 \leq k \leq m$ and distinct $i_1, \dots, i_k \in \{1, \dots, m\}$,

$$\pi_k = \mathbb{P}(Y_{i_1} = 1, \dots, Y_{i_k} = 1) = \widehat{C}_\theta^{\mathrm{Cl}}(\pi, \dots, \pi),$$

where $\widehat{C}_\theta^{\mathrm{Cl}}$ is the survival copula of a k-dimensional Clayton copula $C_\theta^{\mathrm{Cl}}(u_1, \dots, u_k) = (\sum_{j=1}^k u_j^{-\theta} - k + 1)^{-1/\theta}$ with parameter θ that you should determine.

Hint. Show that $\hat{\pi}_k = \mathbb{P}(Y_{i_1} = 0, \dots, Y_{i_k} = 0) = C_\theta^{\mathrm{Cl}}(\hat{\pi}, \dots, \hat{\pi})$ for all $k \in \{1, \dots, m\}$ and use the fact that this implies that $\pi_k = \widehat{C}_\theta^{\mathrm{Cl}}(\pi, \dots, \pi)$ for all $k \in \{1, \dots, m\}$.

References

Adrian, T. and Brunnermeier, M. K. (2016), CoVaR, *American Economic Review*, 106(7), 1705–1741, doi:10.1257/aer.20120555.

Apostol, T. M. (1974), Mathematical Analysis, Addison Wesley.

Arendt, W., Batty, C. J. K., Hieber, M., and Neubrander, F. (2002), Vector-Valued Laplace Transforms and Cauchy Problems, Birkhäuser.

BIS (2016), Minimum capital requirements for market risk, https://www.bis.org/bcbs/publ/d352.pdf (2018-05-18).

BIS (2017), Basel III: Finalising post-crisis reforms, https://www.bis.org/bcbs/publ/d424.pdf (2018-05-18).

Bradley, J.V. (1968), Distribution-Free Statistical Tests, Prentice-Hall.

Brigo, D. and Mercurio, F. (2006), Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit, 2nd, Berlin: Springer.

Brockwell, P. J. and Davis, R. A. (2002), Introduction to Time Series and Forecasting, 2nd, New York: Springer.

Embrechts, P., Puccetti, G., and Rüschendorf, L. (2013), Model uncertainty and VaR aggregation, *Journal of Banking and Finance*, 37(8), 2750–2764.

Genest, C., Rémillard, B., and Beaudoin, D. (2009), Goodness-of-fit tests for copulas: A review and a power study, *Insurance: Mathematics and Economics*, 44, 199–213.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993), On the relation between the expected value and the volatility of the nominal excess return on stocks, *The Journal of Finance*, 48(5), 1779–1801.

Mainik, G. and Schaanning, E. (2014), On dependence consistency of CoVaR and some other systemic risk measures, *Statistics & Risk Modeling*, 31(1), 49–77, doi:10.1515/strm-2013-1164.

Marshall, A. W. and Olkin, I. (1967), A multivariate exponential distribution, *Journal of the American Statistical Association*, 62, 30–44.

McNeil, A. J., Frey, R., and Embrechts, P. (2005), Quantitative Risk Management: Concepts, Techniques and Tools, Princeton: Princeton University Press.

McNeil, A. J., Frey, R., and Embrechts, P. (2015), Quantitative Risk Management: Concepts, Techniques and Tools, 2nd, Princeton: Princeton University Press.

Rosenblatt, M. (1952), Remarks on a Multivariate Transformation, *The Annals of Mathematical Statistics*, 23(3), 470–472.

Scarsini, M. (1984), On measures of concordance, *Stochastica*, 8(3), 201–218.