Course 2023-2024 in Financial Risk Management
Lecture Notes + Tutorial Sessions

Thierry Roncalli*

*Amundi Asset Management[1]

*University of Paris-Saclay

September 2023

[1] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Part 1. Introduction to Financial Risk Management

Part 2. Market Risk

Part 3. Credit Risk

Part 4. Counterparty Credit Risk and Collateral Risk

Part 5. Operational Risk

Part 6. Liquidity Risk

Part 7. Asset Liability Management Risk

Part 8. Model Risk

Part 9. Copulas and Extreme Value Theory

Part 10. Monte Carlo Simulation Methods

Part 11. Stress Testing and Scenario Analysis

Part 12. Credit Scoring Models

Tutorial Session 1

Tutorial Session 2

Tutorial Session 3

Tutorial Session 4

Tutorial Session 5

General information

1. **Overview**
   The objective of this course is to understand the theoretical and practical aspects of risk management
2. **Prerequisites**
   M1 Finance or equivalent
3. **ECTS**
   4
4. **Keywords**
   Finance, Risk Management, Applied Mathematics, Statistics
5. **Hours**
   Lectures: 36h, Training sessions: 15h, HomeWork: 30h
6. **Evaluation**
   There will be a final three-hour exam, which is made up of questions and exercises
7. **Course website**
   `http://www.thierry-roncalli.com/RiskManagement.html`

Objective of the course

The objective of the course is twofold:

1. knowing and understanding the financial regulation (banking and others) and the international standards (especially the Basel Accords)
2. being proficient in risk measurement, including the mathematical tools and risk models

Class schedule

Course sessions

- September 15 (6 hours, AM+PM)
- September 22 (6 hours, AM+PM)
- September 19 (6 hours, AM+PM)
- October 6 (6 hours, AM+PM)
- October 13 (6 hours, AM+PM)
- October 27 (6 hours, AM+PM)

Tutorial sessions

- October 20 (3 hours, AM)
- October 20 (3 hours, PM)
- November 10 (3 hours, AM)
- November 10 (3 hours, PM)
- November 17 (3 hours, PM)

Class times: Fridays 9:00am-12:00pm, 1:00pm–4:00pm, University of Evry, Room 209 IDF

Agenda

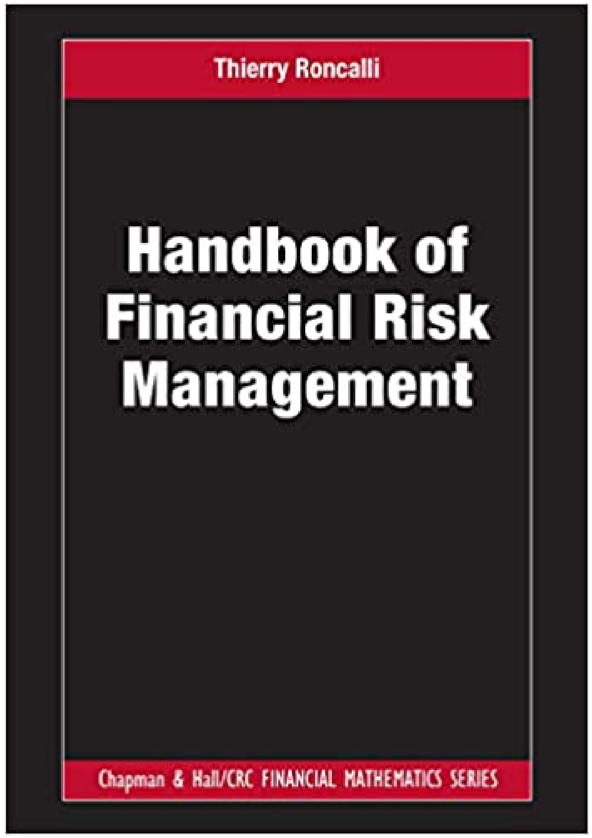
- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Agenda

- Tutorial Session 1: Market Risk
- Tutorial Session 2: Credit Risk
- Tutorial Session 3: Counterparty Credit Risk and Collateral Risk
- Tutorial Session 4: Operational Risk & Asset Liability Management Risk
- Tutorial Session 5: Copulas, EVT & Stress Testing

Textbook

- Roncalli, T. (2020), *Handbook of Financial Risk Management*, Chapman & Hall/CRC Financial Mathematics Series.



Additional materials

- Slides, tutorial exercises and past exams can be downloaded at the following address:

  `http://www.thierry-roncalli.com/RiskManagement.html`

- Solutions of exercises can be found in the companion book, which can be downloaded at the following address:

  `http://www.thierry-roncalli.com/RiskManagementBook.html`

Course 2023-2024 in Financial Risk Management
Lecture 1. Introduction

Thierry Roncalli*

*Amundi Asset Management[2]

*University of Paris-Saclay

September 2023

[2]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- **Lecture 1: Introduction to Financial Risk Management**
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

The development of financial markets

Table: Some financial innovations

1970	Mortgage-backed securities
1971	Equity index funds
1972	Foreign currency futures
1973	Stock options
1979	Over-the-counter currency options
1981	Interest rate swaps
1982	Equity index futures
1983	Equity index options
	Interest rate caps/floors
	Collateralized mortgage obligations
1985	Swaptions
	Asset-backed securities
1987	Path-dependent options (Asian, look-back, etc.)
	Collateralized debt obligations
1994	Credit default swaps
2004	Volatility index futures

The development of financial markets

- Organized markets (on-exchange)
- Over-the-counter markets or OTC markets (off-exchange)

Contract	Futures	Forward	Option	Swap
On-exchange	✓		✓	
Off-exchange		✓	✓	✓

The development of financial markets

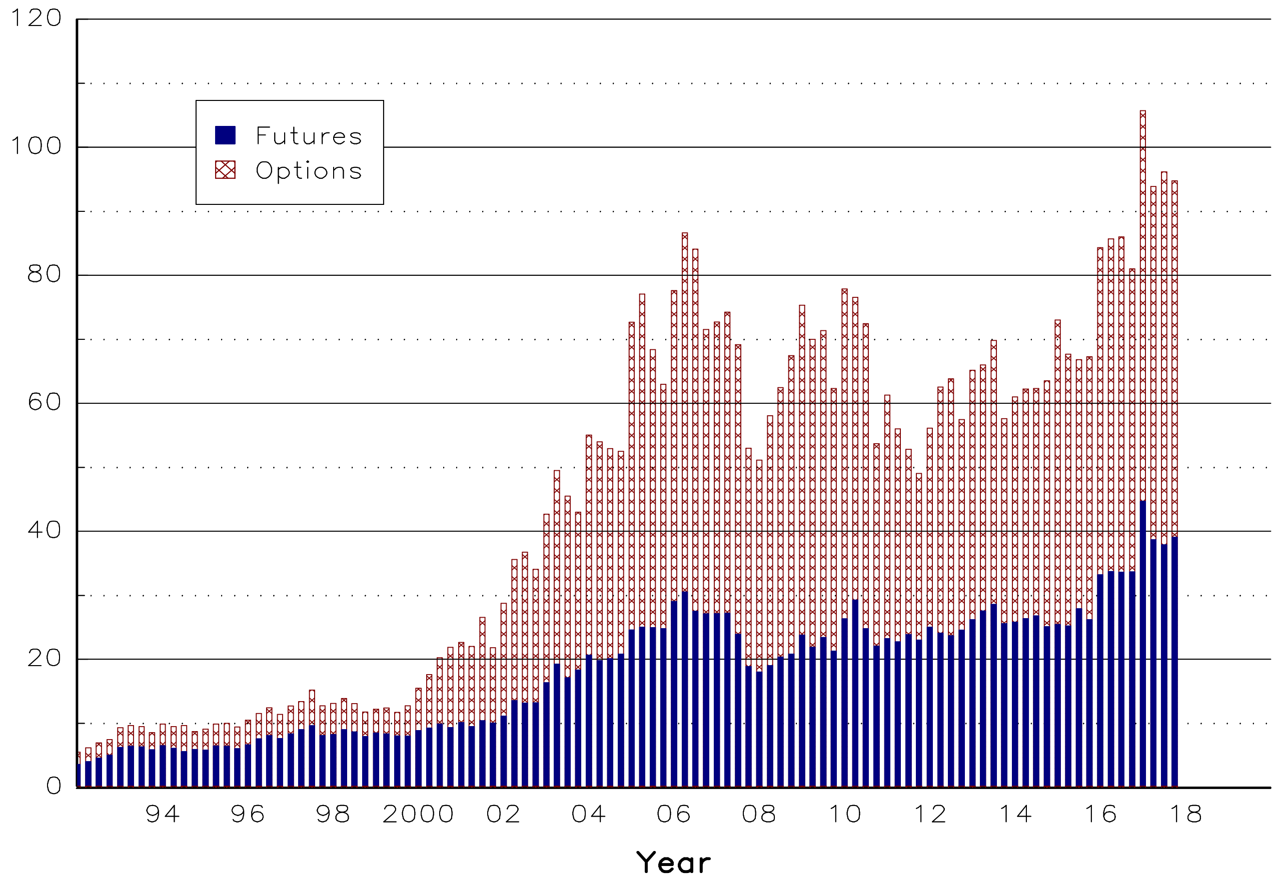


Figure: Notional outstanding amount of exchange-traded derivatives (in $ tn)

Financial crises and systemic risk

Table: Some financial losses

1974	Herstatt Bank: \$620 mn (foreign exchange trading)
1994	Metallgesellschaft: \$1.3 bn (oil futures)
1994	Orange County: \$1.8 bn (reverse repo)
1994	Procter & Gamble: \$160 mn (ratchet swap)
1995	Barings Bank: \$1.3 bn (stock index futures)
1997	Natwest: \$127 mn (swaptions)
1998	LTCM: \$4.6 bn (liquidity crisis)
2001	Dexia Bank: \$270 mn (corporate bonds)
2006	Amaranth Advisors: \$6.5 bn (gaz forward contracts)
2007	Morgan Stanley: \$9.0 bn (credit derivatives)
2008	Société Générale: \$7.2 bn (rogue trading)
2008	Madoff: \$65 bn (fraud)
2011	UBS: \$2.0 bn (rogue trading)
2012	JPMorgan Chase: \$5.8 bn (credit derivatives)

Financial crises and systemic risk

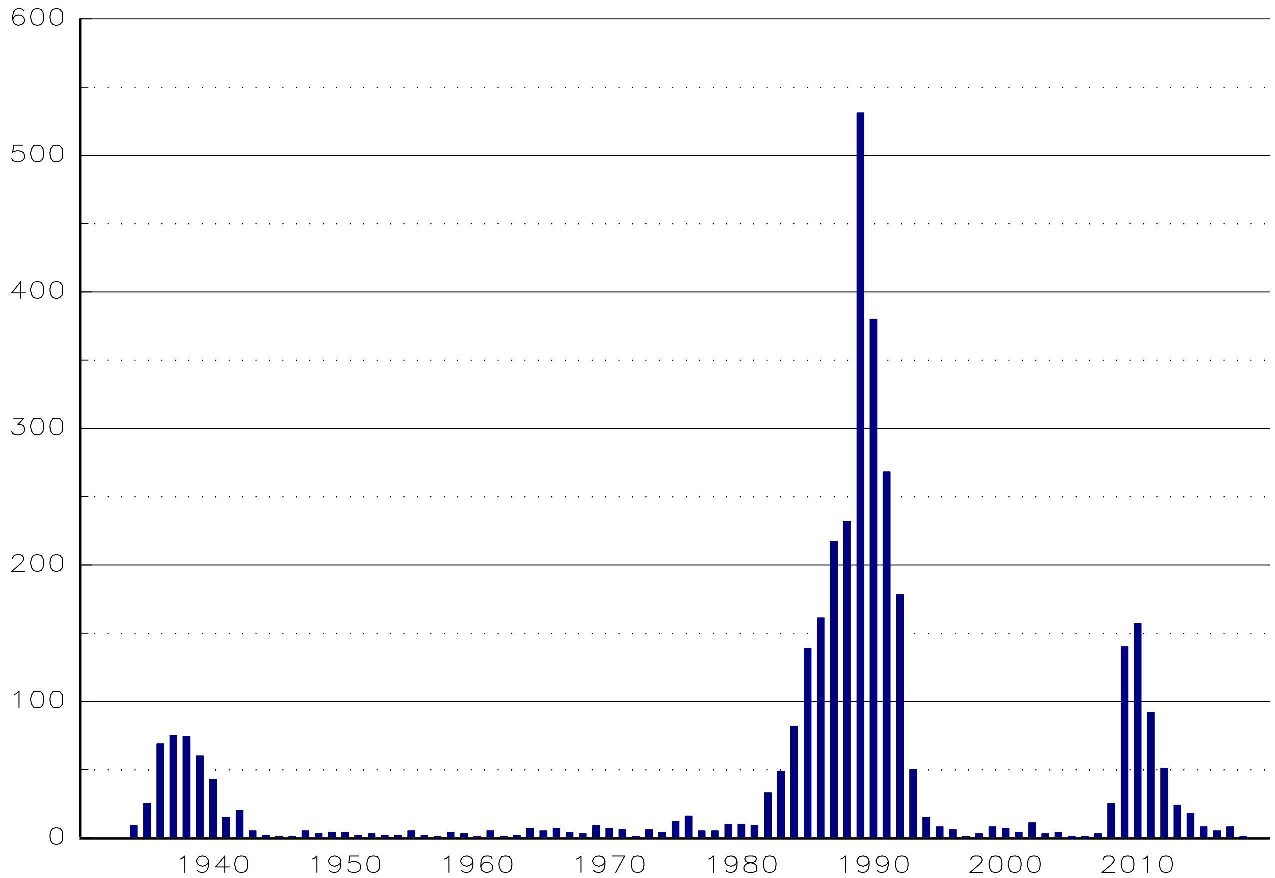


Figure: Number of bank defaults in the US

International authorities

1. The Basel Committee on Banking Supervision (BCBS)
2. The International Association of Insurance Supervisors (IAIS)
3. The International Organization of Securities Commissions (IOSCO)
4. The Financial Stability Board (FSB)

Table: The supervision institutions in finance

	Banks	Insurers	Markets	All sectors
Global	BCBS	IAIS	IOSCO	FSB
EU	EBA/ECB	EIOPA	ESMA	ESFS
US	FDIC/FRB	FIO	SEC	FSOC

Banking regulation

1988 Publication of "*International Convergence of Capital Measurement and Capital Standards*", which is better known as "*The Basel Capital Accord*". This text sets the rules of the Cooke ratio.

1996 Publication of "*Amendment to the Capital Accord to incorporate Market Risks*". This text includes the market risk to compute the Cooke ratio.

2004 Publication of "*International Convergence of Capital Measurement and Capital Standards – A Revisited Framework*". This text establishes the Basel II framework.

2010 Publication of the Basel III framework.

2019 Publication of "*Minimum Capital Requirements for Market Risk*". This is the final version of the Basel III framework for computing the market risk.

Banking regulation

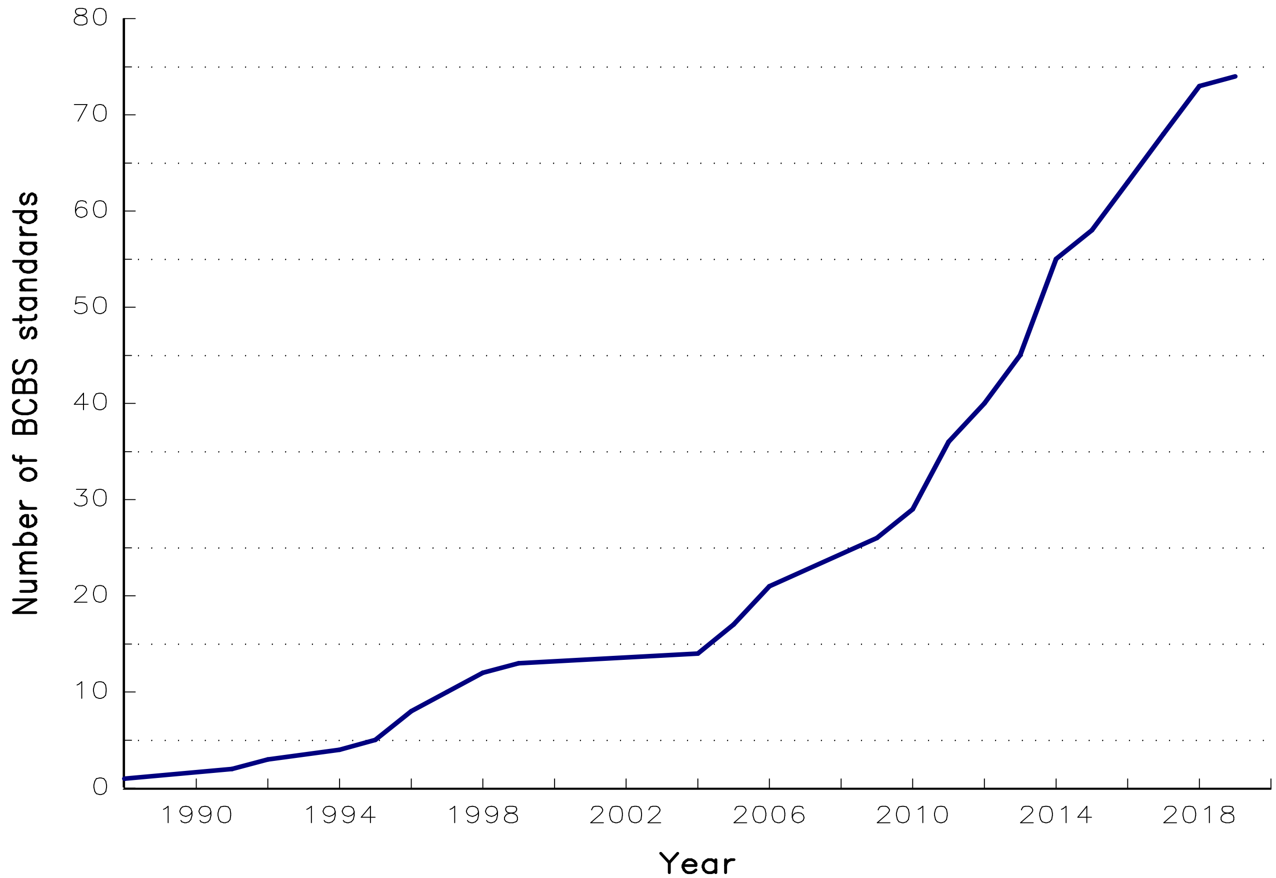


Figure: The huge increase of the number of banking supervision standards

Basel I

- Cooke ratio:

$$\text{Cooke Ratio} = \frac{C}{\mathrm{RWA}}$$

  where C and RWA are the capital and the risk-weighted assets of the bank.

- A risk-weighted asset is simply defined as a bank's asset weighted by its risk score or risk weight (RW):

$$\mathrm{RWA} = \mathrm{EAD} \cdot \mathrm{RW}$$

  where EAD is the exposure at default

$\Rightarrow$ Cooke Ratio $\geq$ 8% (Tier one $\geq$ 4%)

Risk weight

For categories:

1. $\mathrm{RW} = 0\%$
   cash, gold, claims on OECD governments and central banks, claims on governments and central banks outside OECD and denominated in the national currency
2. $\mathrm{RW} = 20\%$
   claims on all banks with a residual maturity lower than one year, longer-term claims on OECD incorporated banks, claims on public-sector entities within the OECD
3. $\mathrm{RW} = 50\%$
   loans secured on residential property
4. $\mathrm{RW} = 100\%$
   others

Computing the RWA

Example

The assets of a bank are composed of \$100 mn of US treasury bonds, \$100 mn of Brazilian government bonds, \$50 mn of residential mortgage, \$300 mn of corporate loans and \$20 mn of revolving credit loans. The bank liability structure includes \$25 mn of common stock and \$13 mn of subordinated debt.

We obtain the following results:

Asset	EAD	RW	RWA
US treasury bonds	100	0%	0
Brazilian Gov. bonds	100	100%	100
Residential mortgage	50	50%	25
Corporate loans	300	100%	300
Revolving credit	20	100%	20
Total			445

and:

$$\text{Cooke Ratio} = \frac{38}{445} = 8.54\%$$

Amendment to incorporate market risks

Two approaches:

- The standardized measurement method (SMM)
- The internal model-based approach[3] (IMA)

⇒ external weights vs internal model (99% value-at-risk for a holding period of 10 trading days)

[3] The use of the internal model-based approach is subject to the approval of the national supervisor.

Value-at-risk (VaR)

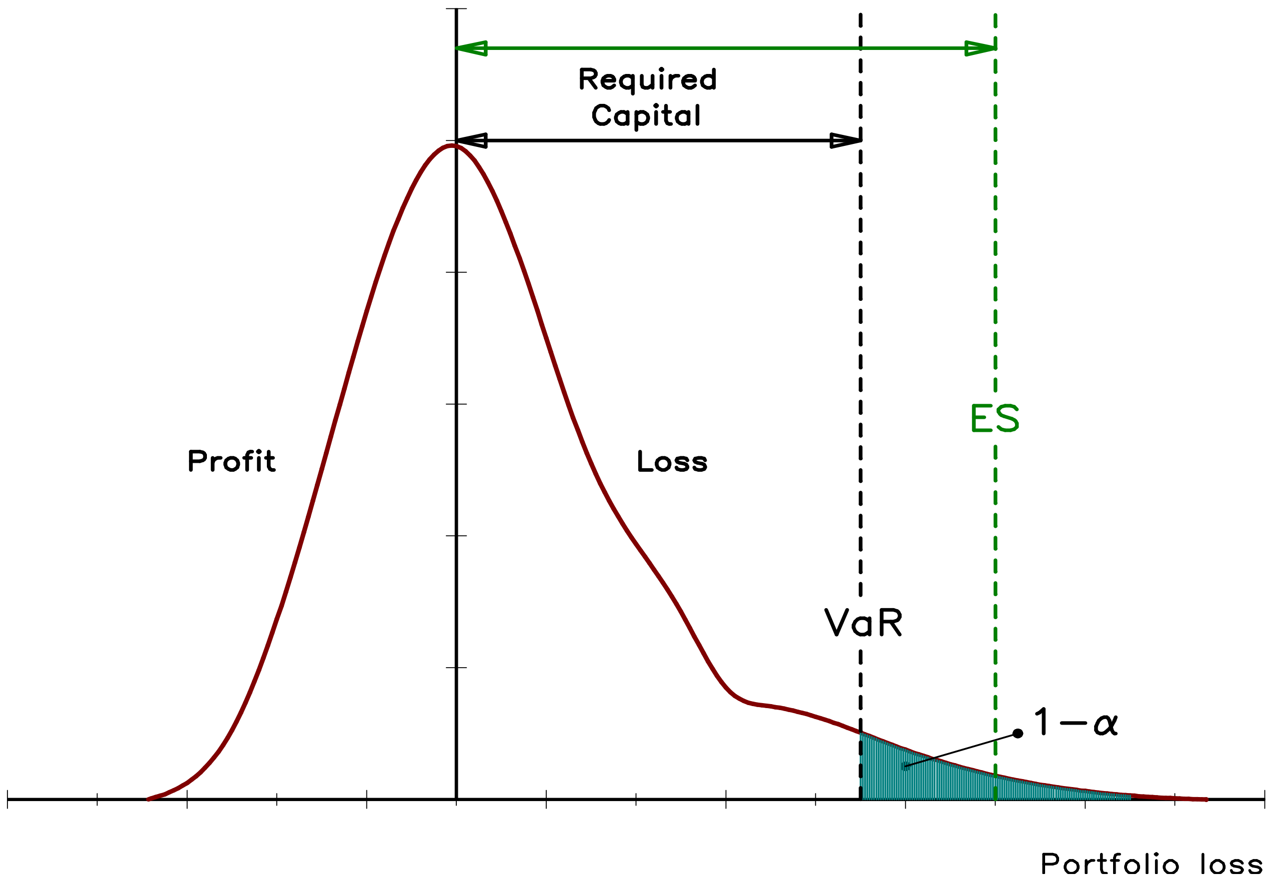


Figure: Probability distribution of the portfolio loss

Impact of market risks on the Cooke ratio

The Cooke ratio becomes:

$$\frac{\mathcal{C}_{\mathrm{Bank}}}{\mathrm{RWA} + 12.5 \times \mathcal{K}_{\mathrm{MR}}} \geq 8\%$$

We deduce that:

$$\mathcal{C}_{\mathrm{Bank}} \geq \underbrace{8\% \times \mathrm{RWA}}_{\mathcal{K}_{\mathrm{CR}}} + \mathcal{K}_{\mathrm{MR}}$$

meaning that $8\% \times \mathrm{RWA}$ can be interpreted as the credit risk capital requirement $\mathcal{K}_{\mathrm{CR}}$, which can be compared to the market risk capital charge $\mathcal{K}_{\mathrm{MR}}$.

Basel II

Table: The three pillars of the Basel II framework

Pillar 1	Pillar 2	Pillar 3
Minimum Capital Requirements	Supervisory Review Process	Market Discipline
Credit risk Market risk Operational risk	Review & reporting Capital above Pillar 1 Supervisory monitoring	Capital structure Capital adequacy Models & parameters Risk management

Basel II

The new Accord consists of three pillars:

1. the first pillar corresponds to *minimum capital requirements*, that is, how to compute the capital charge for credit risk, market risk and operational risk;
2. the second pillar describes the *supervisory review process*; it explains the role of the supervisor and gives the guidelines to compute additional capital charges for specific risks, which are not covered by the first pillar;
3. the *market discipline* establishes the third pillar and details the disclosure of required information regarding the capital structure and the risk exposures of the bank.

Basel II

- Credit risk
  - The standardized approach (SA)
  - The internal ratings-based approach (IRB)
    - Foundation IRB (FIRB or IRB-F)
    - Advanced IRB (AIRB ou IRB-A)
- Market risk
  - The standardized measurement method (SMM)
  - The internal model-based approach (IMA)
- Operational risk
  - The Basic Indicator Approach (BIA)
  - The Standardized Approach (TSA)
  - Advanced Measurement Approaches (AMA)

Basel II

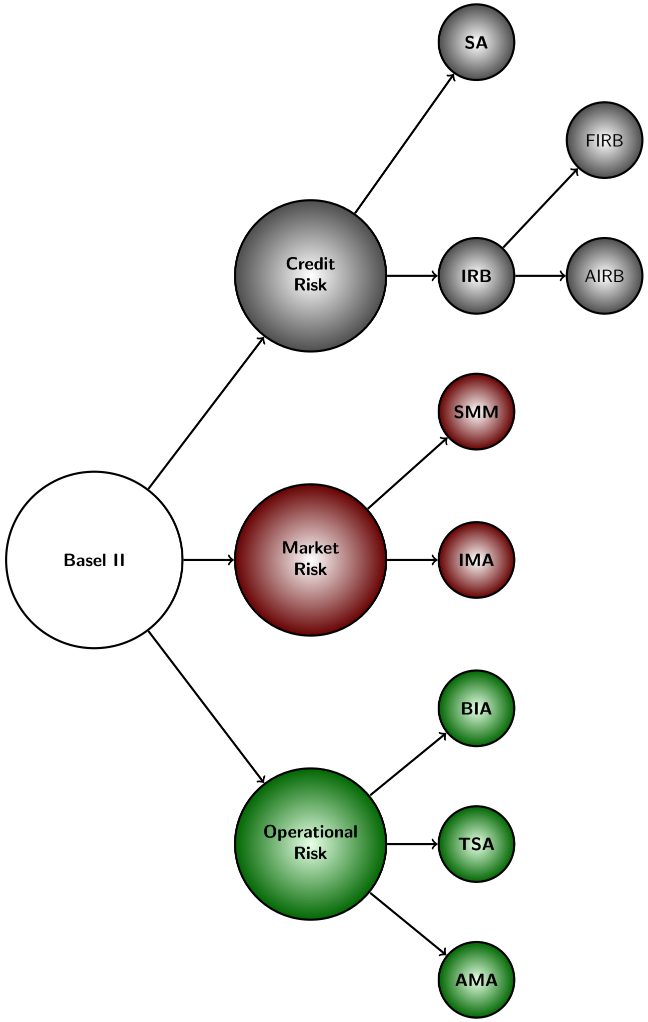


Figure: Minimum capital requirements in the Basel II framework

Basel 2.5

2008 Global Financial Crisis $\Rightarrow$ measures to strengthen the rules governing trading book capital, particularly the market risk associated to securitization and credit-related products:

1. the incremental risk charge (IRC), which is an additional capital charge to capture default risk and migration risk for unsecuritized credit products
2. the stressed value-at-risk requirement (SVaR), which is intended to capture stressed market conditions
3. the comprehensive risk measure (CRM), which is an estimate of risk in the credit correlation trading portfolio (CDS baskets, CDO products, etc.)
4. new standardized charges on securitization exposures, which are not covered by CRM

Basel III

In December 2010, the Basel Committee published a new regulatory framework in order to enhance risk management, increase the stability of the financial markets and improve the banking industry's ability to absorb macro-economic shocks

The Basel III (2010) framework consists of **micro-prudential** and **macro-prudential** regulation measures concerning;

- a new definition of the risk-based capital
- the introduction of a leverage ratio
- the management of the liquidity risk

Basel III also includes (2013-2019):

- Revision of MR, CR, CCR, CVA and OR standards
- Interest Rate Risk in the Banking Book (IRRBB)

Basel III

Table: Basel III capital requirements

Capital ratio	2013	2014	2015	2016	2017	2018	2019
CET1	3.5%	4.0%	4.5%				**4.5%**
CB				0.625%	1.25%	1.875%	**2.5%**
CET1 + CB	3.5%	4.0%	4.5%	5.125%	5.75%	6.375%	**7.0%**
Tier 1	4.5%	5.5%	6.0%				**6.0%**
Total	8.0%						**8.0%**
Total + CB	8.0%			8.625%	9.25%	9.875%	**10.5%**
CCB				**0% – 2.5%**			

- CET1: Common Equity Tier 1
- AT1: Additional Tier 1
- T1: Tier 1
- T2: Tier 2
- CB: Capital Conservation Buffer
- CCB: Countercyclical Conservation Buffer (**macro-prudential** measure)

Basel III

- Credit Valuation Adjustment (CVA)
- Leverage ratio (**macro-prudential** measure) to prevent the build-up of excessive on- and off-balance sheet:

$$\text{Leverage ratio} = \frac{\text{Tier 1 capital}}{\text{Total exposures}} \geq 3\%$$

where the total exposures is the sum of on-balance sheet exposures, derivative exposures and some adjustments concerning off-balance sheet items

Basel III

- Liquidity Coverage Ratio (LCR)
  The objective of the LCR is to promote short-term resilience of the bank's liquidity risk profile:

$$\mathrm{LCR} = \frac{\mathrm{HQLA}}{\mathrm{Total\ net\ cash\ outflows}} \geq 100\%$$

  where HQLA is the stock of high quality liquid assets and the denominator is the total net cash outflows over the next 30 calendar days

- Net Stable Funding Ratio (NSFR)
  NSFR is designed in order to promote long-term resilience of the bank's liquidity profile:

$$\mathrm{NSFR} = \frac{\text{Available amount of stable funding}}{\text{Required amount of stable funding}} \geq 100\%$$

  ASF and RSF are calculated for the next year

Basel III

Basel III also includes new standards (the Basel IV package):

- Credit Risk: revision to SA and IRB approaches
- Market Risk: SMM is replaced by SA-TB, IMA is revisited, VaR is replaced by ES (expected shortfall), etc.
- CVA $\Rightarrow$ SA-CVA and BA-CVA
- Operational Risk: BIA, TSA and AMA are replaced by SMA (Standardized Measurement Approach)
- Introduction of capital floors (with respect to SA)

Insurance regulation

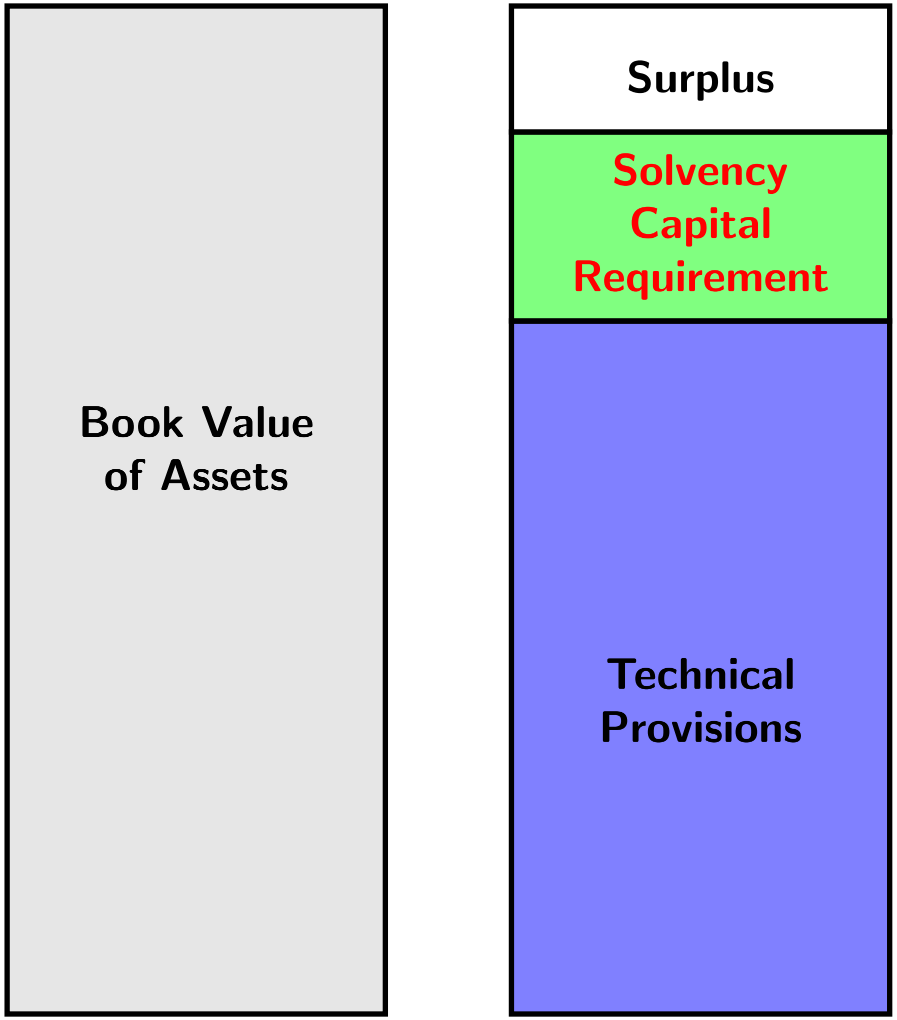


Figure: Solvency I capital requirement

Insurance regulation

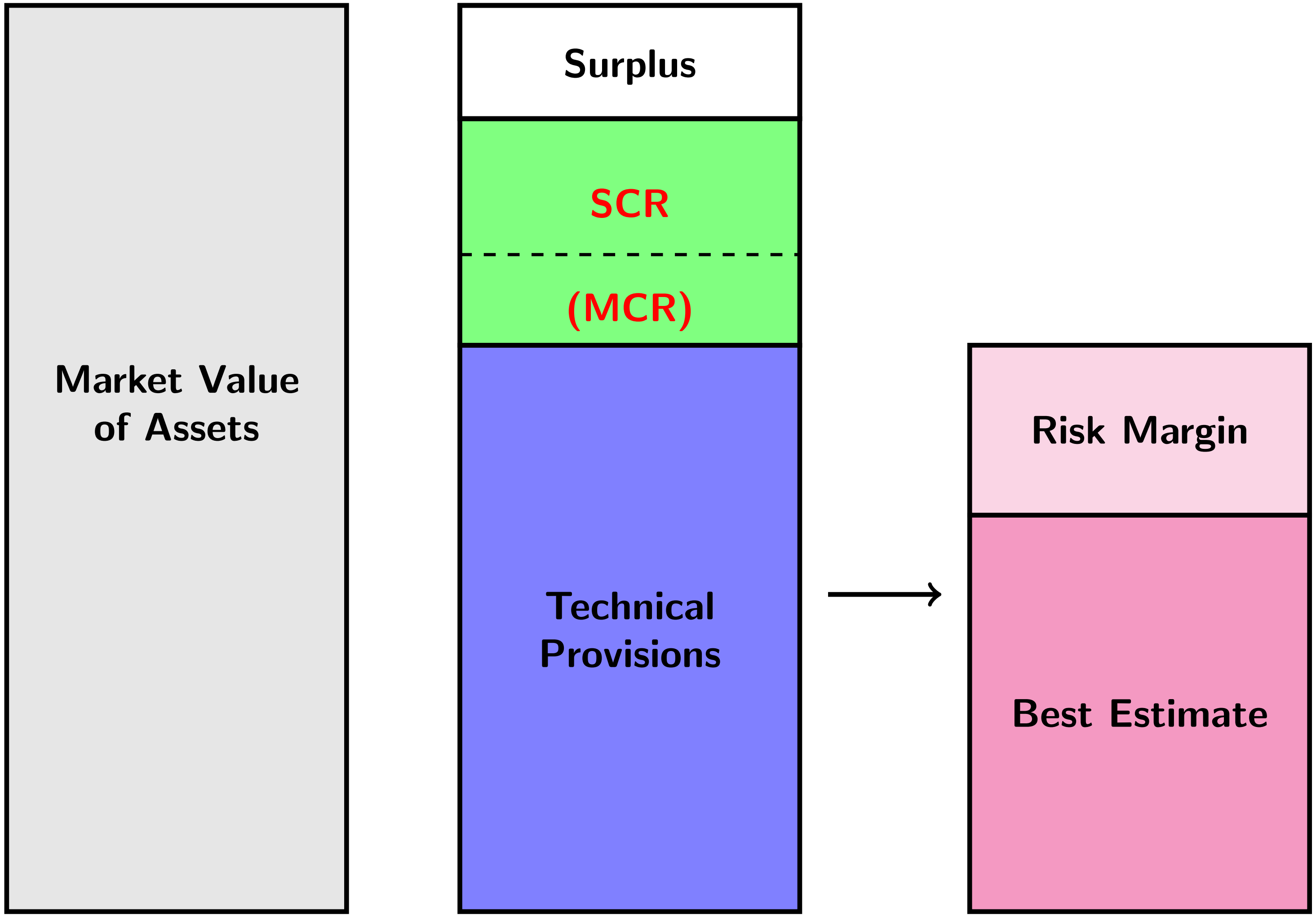


Figure: Solvency II capital requirement

Insurance regulation

Risk components:

1. Underwriting risk (non-life, life, health, etc.)
2. market risk,
3. Default risk
4. Counterparty credit risk

In the case of the standard formula method, the SCR of the insurer is equal to:

$$\mathrm{SCR} = \sqrt{\sum_{i,j}^{m} \rho_{i,j} \cdot \mathrm{SCR}_i \cdot \mathrm{SCR}_j} + \mathrm{SCR}_{\mathrm{OR}}$$

where SCR_i is the SCR of the risk module i, $\mathrm{SCR}_{\mathrm{OR}}$ is the SCR associated to the operational risk and $\rho_{i,j}$ is the correlation factor between risk modules i and j.

Insurance regulation

The solvency ratio is then defined as:

$$\text{Solvency Ratio} = \frac{C}{\mathrm{SCR}}$$

where C is the capital. This solvency ratio must be larger than 33% for tier 1 and 100% for the total own funds.

Market regulation

Europe

- 2007: MiFID (Markets in Financial Instruments Directive)
- 2012: EMIR (European Market Infrastructure Regulation)
- 2014: MiFID2, MiFIR (Regulation in Markets in Financial Instruments) and PRIIPS (Packaged Retail and Insurance-based Investment Products)

US

- 1930s: Securities Act, Securities Exchange Act, Trust Indenture Act, Investment Company Act, Investment Advisers Act
- Securities and Exchange Commission (SEC)
- Commodity Futures Trading Commission (CFTC)
- 2010: Dodd-Frank Wall Street Reform and Consumer Protection Act
- Financial Stability Oversight Council (FSOC)

Systemic risk

- 2009: Creation of the Financial Stability Board (FSB)
- Systemically Important Financial Institutions (SIFIs)
- A SIFI can be global (G-SIFI) or domestic (D-SIFI)
- Three categories:
  1. G-SIBs correspond to global systemically important banks
  2. G-SIIs designate global systemically important insurers
  3. The third category corresponds to non-bank non-insurer global systemically important financial institutions (or NBNI G-SIFIs)

Course 2023-2024 in Financial Risk Management
Lecture 2. Market Risk

Thierry Roncalli*

*Amundi Asset Management[4]

*University of Paris-Saclay

September 2023

[4]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction to Financial Risk Management
- **Lecture 2: Market Risk**
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Most important dates

- 19 October 1987: Stock markets crashed and the Dow Jones Industrial Average index dropped by more than 20% in the day
- 1988: Publication of the Basel I Accord
- 1990s: Japanese asset price bubble
- 1994: Bond market massacre
- October 1994: Publication of *RiskMetrics* by J.P. Morgan
- January 1996: Amendment to incorporate market risks (Basel I)
- 2004: Measuring market risks is the same in Basel II
- 2008: Global Financial Crisis (GFC)
- 2009: Basel 2.5
- January 2019: Revision of market risk in Basel III (also known as the **fundamental review of the trading book** or FRTB)

Definition

According to the Basel Committee, market risk is defined as "*the risk of losses (in on- and off-balance sheet positions) arising from movements in market prices. The risks subject to market risk capital requirements include but are not limited to:*

- *default risk, interest rate risk, credit spread risk, equity risk, foreign exchange (FX) risk and commodities risk for trading book instruments*;
- *FX risk and commodities risk for banking book instruments*."

Portfolio	Fixed Income	Equity	Currency	Commodity	Credit
Trading	✓	✓	✓	✓	✓
Banking			✓	✓	

$\Rightarrow$ trading book $\neq$ banking book

The Basel I/II framework

To compute the capital charge, banks have the choice between two approaches:

1. the standardized measurement method (SMM)
2. the internal model-based approach (IMA)

$\Rightarrow$ Banks quickly realized that they can sharply reduce their capital requirements by adopting internal models

Standardized measurement method (SMM)

Five main risk categories:

1. Interest rate risk
2. Equity risk
3. Currency risk
4. Commodity risk
5. Price risk on options and derivatives

For each category, a capital charge is computed to cover:

- the general market risk
- the specific risk

Standardized measurement method (SMM)

The capital charge $\mathcal{K}$ is equal to the risk exposure E times the capital charge weight K:

$$\mathcal{K} = E \cdot K$$

- For the specific risk, the risk exposure corresponds to the notional of the instrument, whether it is a long or a short position
- For the general market risk, long and short positions on different instruments can be offset

The case of equity risk

- The capital charge for specific risk is 4% if the portfolio is liquid and well-diversified and 8% otherwise
- For the general market risk, the risk weight is equal to 8% and applies to the net exposure

Remark

Under Basel 2.5, the capital charge for specific risk is set to 8% whatever the liquidity of the portfolio

The case of equity risk

Example

We consider a $100 mn short exposure on the S&P 500 index futures contract and a $60 mn long exposure on the Apple stock.

The capital charge for specific risk is[5]:

$$\mathcal{K}^{\text{Specific}} = 100 \times 4\% + 60 \times 8\% = 4 + 4.8 = 8.8$$

The net exposure is −$40 mn. We deduce that the capital charge for the general market risk is:

$$\mathcal{K}^{\text{General}} = |-40| \times 8\% = 3.2$$

It follows that the total capital charge for this equity portfolio is $12 mn.

[5] We assume that the S&P 500 index is liquid and well-diversified, whereas the exposure on the Apple stock is not diversified.

The case of interest rate risk (specific risk)

- For government instruments, the capital charge weights are:

Rating	AAA to AA−	A+ to BBB−			BB+ to B−	Below B−	NR
Maturity		0−6M	6M−2Y	2Y+			
K	0%	0.25%	1.00%	1.60%	8%	12%	8%

- In the case of other instruments (PSE, banks and corporates), the capital charge weights are:

Rating	AAA to BBB−			BB+ to BB−	Below BB−	NR
Maturity	0−6M	6M−2Y	2Y+			
K	0.25%	1.00%	1.60%	8%	12%	8%

The case of interest rate risk (specific risk)

Example

We consider a trading portfolio with the following exposures: a long position of \$50 mn on Euro-Bund futures, a short position of \$100 mn on three-month T-Bills and a long position of \$10 mn on an investment grade (IG) corporate bond with a three-year residual maturity.

$\Rightarrow$ Why the capital charge for specific risk is equal to \$0, \$0 and \$160 000?

The case of interest rate risk (general market risk)

Two methods:

- Maturity approach
- Duration approach (price sensitivity with respect to a change in yield)

Internal model-based approach

The use of an internal model is conditional upon the approval of the supervisory authority:

- **Qualitative criteria**
  - Independent risk control unit
  - Daily reports
  - Daily risk management
  - Etc.
- **Quantitative criteria**
  - The value-at-risk (VaR) is computed on a daily basis with a 99% confidence level. The minimum holding period of the VaR is 10 trading days. If the bank computes a VaR with a shorter holding period, it can use the square-root-of-time rule
  - Relevant risk factors
  - Sample period: at least one year
  - The value of the multiplication factor depends on the quality of the internal model with a range between 3 and 4. The quality of the internal model is related to its ex-post performance measured by the backtesting procedure
  - **Stress testing & Backtesting**

The square-root-of-time rule

The holding period to define the capital is 10 trading days. For that, banks can compute the one-day VaR and converts it to a ten-day VaR:

$$\mathrm{VaR}_{\alpha}\left(w;\text{ten days}\right)=\sqrt{10}\times\mathrm{VaR}_{\alpha}\left(w;\text{one day}\right)$$

Required capital

The required capital at time t is equal to:

$$\mathcal{K}_t = \max\left(\mathrm{VaR}_{t-1}, (3+\xi)\cdot\frac{1}{60}\sum_{i=1}^{60}\mathrm{VaR}_{t-i}\right)$$

where VaR_t is the 10-day value-at-risk calculated at time t and ξ is the penalty coefficient ($0 \leq \xi \leq 1$)

Required capital

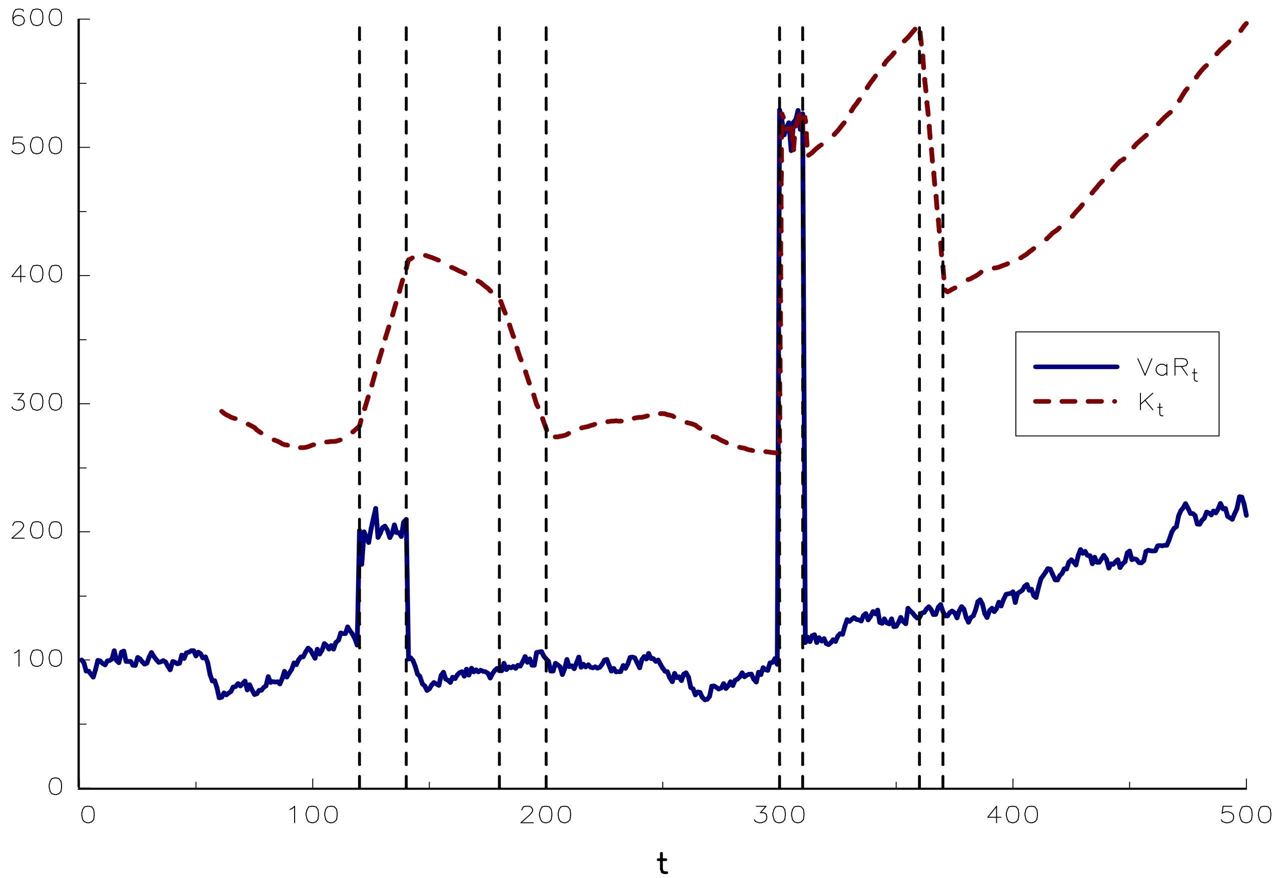


Figure: Calculation of the required capital with the VaR

Backtesting

Definition

Backtesting consists of verifying that the internal model is consistent with a 99% confidence level

⇒ For instance, we expect that the realized loss exceeds the VaR figure once every 100 observations on average

Table: Value of the penalty coefficient ξ for a sample of 250 observations

Zone	Number of exceptions	ξ
Green	0 – 4	0.00
Yellow	5	0.40
	6	0.50
	7	0.65
	8	0.75
	9	0.85
Red	10+	1.00

Statistical approach of backtesting

We note w the portfolio, $\mathrm{VaR}_\alpha(w;h)$ the value-at-risk calculated at time $t-1$, and $L_t(w)$ the daily loss at time t:

$$L_t(w) = -\Pi_t(w) = \mathrm{MtM}_{t-1} - \mathrm{MtM}_t$$

By definition, we have:

$$\Pr\{L_t(w) \geq \mathrm{VaR}_\alpha(w;h)\} = 1-\alpha$$

Let e_t be the random variable which is equal to 1 if there is an exception and 0 otherwise. e_t is a Bernoulli random variable with parameter p:

$$p = \Pr\{e_t = 1\} = \Pr\{L_t(w) \geq \mathrm{VaR}_\alpha(w;h)\} = 1-\alpha$$

Let $N_e(t_1;t_2) = \sum_{t=t_1}^{t_2} e_t$ be the number of exceptions for the period $[t_1, t_2]$. We assume that the exceptions are independent across time.

Main result

$N_e(t_1;t_2)$ is a binomial random variable $\mathcal{B}(n; 1-\alpha)$

Statistical approach of backtesting

Table: Probability distribution (in %) of the number of exceptions ($n = 250$ trading days)

	$\alpha = 99\%$		$\alpha = 98\%$	
m	$\Pr\{N_e = m\}$	$\Pr\{N_e \leq m\}$	$\Pr\{N_e = m\}$	$\Pr\{N_e \leq m\}$
0	8.106	8.106	0.640	0.640
1	20.469	28.575	3.268	3.908
2	25.742	54.317	8.303	12.211
3	21.495	75.812	14.008	26.219
4	13.407	89.219	17.653	43.872
5	6.663	95.882	17.725	61.597
6	2.748	98.630	14.771	76.367
7	0.968	99.597	10.507	86.875
8	0.297	99.894	6.514	93.388
9	0.081	99.975	3.574	96.963
10	0.020	99.995	1.758	98.720

Statistical approach of backtesting

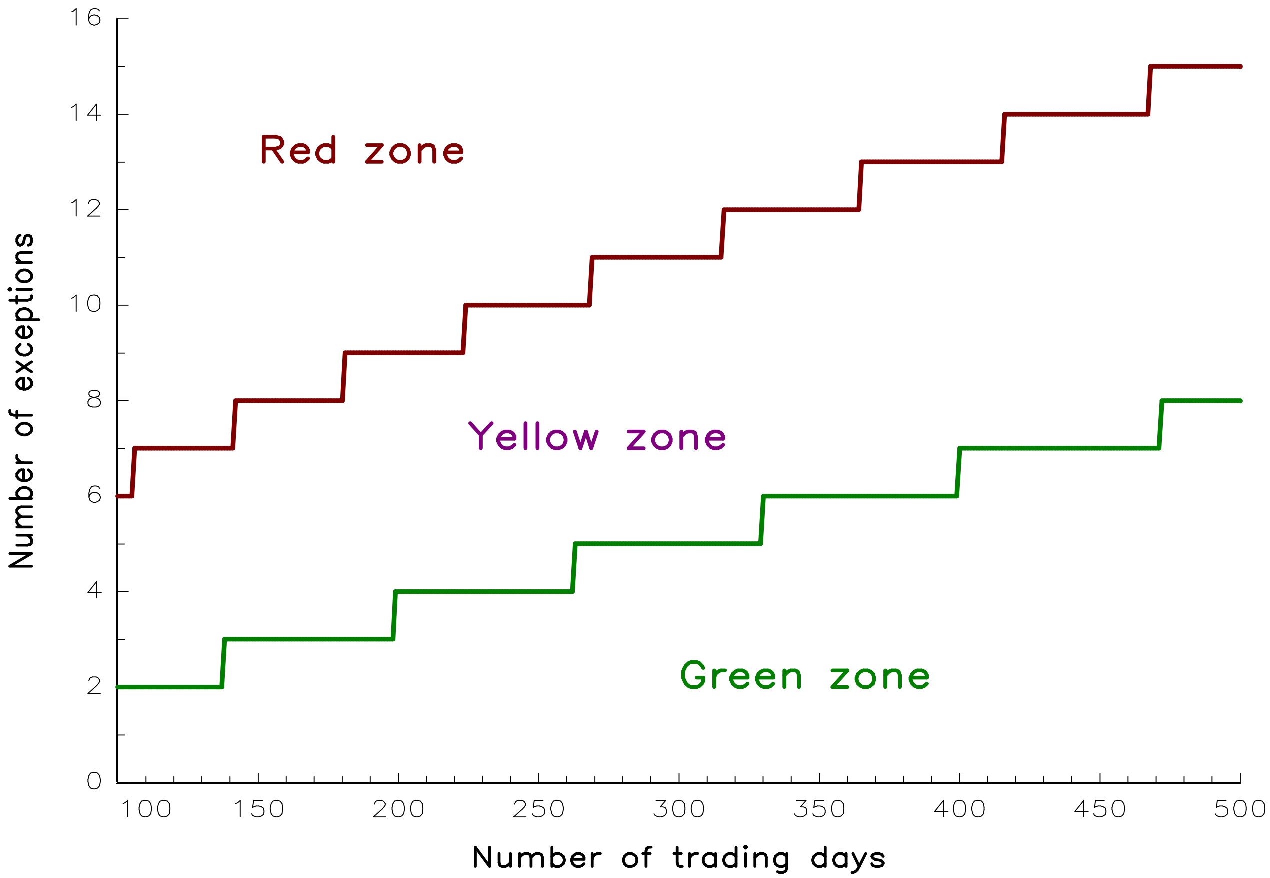


Figure: Color zones of the backtesting procedure ($\alpha = 99\%$)

The Basel 2.5 framework

The required capital becomes:

$$\mathcal{K}_t = \mathcal{K}_t^{\mathrm{VaR}} + \mathcal{K}_t^{\mathrm{SVaR}} + \mathcal{K}_t^{\mathrm{SRC}} + \mathcal{K}_t^{\mathrm{IRC}} + \mathcal{K}_t^{\mathrm{CRM}}$$

where $\mathcal{K}_t^{\mathrm{VaR}}$ is the VaR capital and $\mathcal{K}_t^{\mathrm{SRC}}$ (Basel II), and:

- $\mathcal{K}_t^{\mathrm{SVaR}}$ is the **Stressed VaR**
- $\mathcal{K}_t^{\mathrm{IRC}}$ is the **incremental risk charge** (IRC), which measures the impact of rating migrations and defaults
- $\mathcal{K}_t^{\mathrm{CRM}}$ is the **comprehensive risk measure** (CRM), which corresponds to a supplementary capital charge for credit exotic trading portfolios

The stressed VaR

Definition

The stressed VaR has the same characteristics than the traditional VaR (99% confidence level and 10-day holing period), but the model inputs are "*calibrated to historical data from a continuous 12-month period of significant financial stress relevant to the bank's portfolio*".

⇒ This implies that the historical period to compute the SVaR is completely different than the historical period to compute the VaR[6]

[6]For instance, a typical period is the 2008 year which both combines the subprime mortgage crisis and the Lehman Brothers bankruptcy

The Basel III framework

Banks have the choice between two approaches for computing the capital charge:

1. a standardized method (SA-TB[7])
2. an internal model-based approach (IMA)

⇒ SMM is replaced by SA-TB and IMA is revisited

Remark

Contrary to the previous framework, the SA-TB method is very important even if banks calculate the capital charge with the IMA method. Indeed, the bank must implement SA-TB in order to meet the output (or capital) floor requirement, which is set at 72.5% in January 2027:

$$\mathcal{K}_t = \max\left(\mathcal{K}_t^{IMA}, 72.5\% \times \mathcal{K}_t^{SA\text{-}TB}\right)$$

[7] TB means trading book

SA-TB

The standardized capital charge is the sum of three components:

1. sensitivity-based capital requirement
2. the default risk capital (DRC)
3. the residual risk add-on (RRAO)

Some comments:

- The first component must be viewed as the pure market risk and is the equivalent of the capital charge for the general market risk
- The second component captures the jump-to-default risk (JTD) and replaces the specific risk
- The last component captures specific risks that are difficult to measure in practice

Sensitivity-based capital requirement

We have:

$$\mathcal{K} = \mathcal{K}^{\text{Delta}} + \mathcal{K}^{\text{Vega}} + \mathcal{K}^{\text{Curvature}}$$

$\Rightarrow$ a capital charge for delta, vega and curvature risks

7 risk classes:

1. General interest rate risk (GIRR)
2. Credit spread risk(CSR) on non-securitization products
3. Credit spread risk(CSR) on non-correlation trading portfolio (non-CTP)
4. Credit spread risk(CSR) on correlation trading portfolio (CTP)
5. Equity risk
6. Commodity risk
7. Foreign exchange risk

Delta and vega risk components

- We first begin to calculate the weighted sensitivity of each risk factor $\mathcal{F}_j$:
$$\mathrm{WS}_j = S_j \cdot \mathrm{RW}_j$$
where S_j and RW_j are the net sensitivity of the portfolio with respect to the risk factor and the risk weight of $\mathcal{F}_j$
- Second, we calculate the capital requirement for the risk bucket $\mathcal{B}_k$:
$$\mathcal{K}_{\mathcal{B}_k} = \sqrt{\max\left(\sum_j \mathrm{WS}_j^2 + \sum_{j' \neq j} \rho_{j,j'} \,\mathrm{WS}_j \,\mathrm{WS}_{j'}, 0\right)}$$
where $\mathcal{F}_j \in \mathcal{B}_k$.
- Finally, we aggregate the different buckets for a given risk class:
$$\mathcal{K}^{\mathrm{Delta/Vega}} = \sqrt{\sum_k \mathcal{K}_{\mathcal{B}_k}^2 + \sum_{k' \neq k} \gamma_{k,k'} \,\mathrm{WS}_{\mathcal{B}_k} \,\mathrm{WS}_{\mathcal{B}_{k'}}}$$
where $\mathrm{WS}_{\mathcal{B}_k} = \sum_{j \in \mathcal{B}_k} \mathrm{WS}_j$ is the weighted sensitivity of the bucket $\mathcal{B}_k$.

Delta and vega risk components

The capital requirement for delta and vega risks can be viewed as a Gaussian risk measure with the following parameters:

1. the sensitivities S_j of the risk factors that are calculated by the bank;
2. the risk weights RW_j of the risk factors;
3. the correlation $\rho_{j,j'}$ between risk factors within a bucket;
4. the correlation $\gamma_{k,k'}$ between the risk buckets.

Curvature risk component

The curvature risk uses a similar methodology, but it is based on two adverse scenarios: (1) the risk factor is shocked upward and (2) the risk factor is shocked downward

The curvature risk is close to the gamma risk that we encounter in the theory of options

Practical computation of dela, vega and curvature risks

Three steps:

1. defining the risk factors
2. calculating the sensitivities
3. calculating the risk-weighted sensitivities WS_j

Defining the risk factors

⇒ The Basel Committee gives a very precise list of risk factors by asset classes

For instance, the equity delta risk factors are the equity spot prices and the equity repo rates, the equity vega risk factors are the implied volatilities of options, and the equity curvature risk factors are the equity spot prices

In the case of the interest rate risk class (GIRR), the risk factors include the yield curve[8], a flat curve of market-implied inflation rates for each currency and some cross-currency basis risks

[8] The risk factors correspond to the following tenors of the yield curve: 3M, 6M, 1Y, 2Y, 3Y, 5Y, 10Y, 15Y, 20Y and 30Y.

Calculating the sensitivities

The equity delta sensitivity of the instrument i with respect to the equity risk factor $\mathcal{F}_j$ is given by:

$$S_{i,j} = \mathbf{\Delta}_i \left(\mathcal{F}_j\right) \cdot \mathcal{F}_j$$

where $\mathbf{\Delta}_i \left(\mathcal{F}_j\right)$ measures the (discrete) delta of the instrument i by shocking the equity risk factor $\mathcal{F}_j$ by 1%:

$$S_{i,j} = \frac{P_i \left(1.01 \cdot \mathcal{F}_j\right) - P_i \left(\mathcal{F}_j\right)}{1.01 \cdot \mathcal{F}_j - \mathcal{F}_j} \cdot \mathcal{F}_j = \frac{P_i \left(1.01 \cdot \mathcal{F}_j\right) - P_i \left(\mathcal{F}_j\right)}{0.01}$$

Remark

- *If the instrument corresponds to a stock, the sensitivity is exactly the price of this stock when the risk factor is the stock price, and zero otherwise*
- *If the instrument corresponds to an European option on this stock, the sensitivity is the traditional delta of the option times the stock price*

Calculating the sensitivities

For the vega sensitivity, we have:

$$S_{i,j} = \boldsymbol{\upsilon}_i \left(\mathcal{F}_j\right) \cdot \mathcal{F}_j$$

where $\mathcal{F}_j$ is the implied volatility and $\boldsymbol{\upsilon}_i \left(\mathcal{F}_j\right)$ is the vega of the instrument

Calculating the risk-weighted sensitivities

We use the figures given in BCBS (2019) for the risk weight RW_j, the correlation $\rho_{j,j'}$ and the correlation $\gamma_{k,k'}$

Internal model-based approach

A trading desk is "*an unambiguously defined group of traders or trading accounts that implements a well-defined business strategy operating within a clear risk management structure*".

$\Rightarrow$ Internal models are implemented at the **trading desk** level, meaning that some trading desks are approved for the use of internal models, while other trading desks must use the SA-TB approach

Capital requirement for modellable risk factors

Main differences with Basel I/II

The value-at-risk at the 99% confidence level is replaced by the expected shortfall at the 97.5% confidence level. Moreover, the 10-day holding period is not valid for all instruments

Expected shortfall

The expected shortfall is the average loss beyond the value-at-risk

Capital requirement for modellable risk factors

Impact of the liquidity

$$\mathrm{ES}_\alpha(w) = \sqrt{\sum_{k=1}^{5}\left(\mathrm{ES}_\alpha(w; h_k)\sqrt{\frac{h_k - h_{k-1}}{h_1}}\right)^2}$$

- $\mathrm{ES}_\alpha(w; h_1)$ is the expected shortfall of the portfolio w at horizon 10 days by considering all risk factors
- $\mathrm{ES}_\alpha(w; h_k)$ is the expected shortfall of the portfolio w at horizon h_k days by considering the risk factors $\mathcal{F}_j$ that belongs to the liquidity class k
- h_k is the horizon of the liquidity class k, which is given below:

Liquidity class k	1	2	3	4	5
Liquidity horizon h_k	10	20	40	60	120

Capital requirement for modellable risk factors

Liquidity buckets

1. Interest rates (specified currencies and domestic currency of the bank), equity prices (large caps), FX rates (specified currency pairs).
2. Interest rates (unspecified currencies), equity prices (small caps) and volatilities (large caps), FX rates (currency pairs), credit spreads (IG sovereigns), commodity prices (energy, carbon emissions, precious metals, non-ferrous metals).
3. FX rates (other types), FX volatilities, credit spreads (IG corporates and HY sovereigns).
4. Interest rates (other types), IR volatility, equity prices (other types) and volatilities (small caps), credit spreads (HY corporates), commodity prices (other types) and volatilities (energy, carbon emissions, precious metals, non-ferrous metals).
5. Credit spreads (other types) and credit spread volatilities, commodity volatilities and prices (other types).

Capital requirement for modellable risk factors

How to calculate the expected shortfall for a period of stress?

$$\mathrm{ES}_{\alpha}(w;h) = \mathrm{ES}_{\alpha}^{(\mathrm{reduced,stress})}(w;h) \cdot \min\left(\frac{\mathrm{ES}_{\alpha}^{(\mathrm{full,current})}(w;h)}{\mathrm{ES}_{\alpha}^{(\mathrm{reduced,current})}(w;h)}, 1\right)$$

where $\mathrm{ES}_{\alpha}^{(\mathrm{full,current})}$ is the expected shortfall based on the current period with the full set of risk factors, $\mathrm{ES}_{\alpha}^{(\mathrm{reduced,current})}$ is the expected shortfall based on the current period with a restricted set of risk factors and $\mathrm{ES}_{\alpha}^{(\mathrm{reduced,stress})}$ is the expected shortfall based on the stress period with the restricted set of risk factors

Remark

The previous formula assumes that there is a proportionality factor between the full set and the restricted set of risk factors:

$$\frac{\mathrm{ES}_{\alpha}^{(\mathrm{full,stress})}(w;h)}{\mathrm{ES}_{\alpha}^{(\mathrm{full,current})}(w;h)} \approx \frac{\mathrm{ES}_{\alpha}^{(\mathrm{reduced,stress})}(w;h)}{\mathrm{ES}_{\alpha}^{(\mathrm{reduced,current})}(w;h)}$$

Capital requirement for modellable risk factors

Example

In the table below, we have calculated the 10-day expected shortfall for a given portfolio:

Set of risk factors	Period	Liquidity class 1	2	3	4	5
Full	Current	100	75	34	12	6
Reduced	Current	88	63	30	7	5
Reduced	Stress	112	83	47	9	7

Capital requirement for modellable risk factors

Table: Scaled expected shortfall

k	Sc_k	Full Current	Reduced Current	Reduced Stress	Full/Stress (not scaled)	Full Stress
1	1	100.00	88.00	112.00	127.27	127.27
2	1	75.00	63.00	83.00	98.81	98.81
3	$\sqrt{2}$	48.08	42.43	66.47	53.27	75.33
4	$\sqrt{2}$	16.97	9.90	12.73	15.43	21.82
5	$\sqrt{6}$	14.70	12.25	17.15	8.40	20.58
Total		135.80	117.31	155.91		180.38

The scaling factor is equal to $Sc_k = \sqrt{(h_k - h_{k-1}) / h_1}$, the scaled expected shortfall is equal to $\mathrm{ES}_\alpha^\star (w; h_k) = Sc_k \cdot \mathrm{ES}_\alpha (w; h_k)$ and the total expected shortfall is given by $\mathrm{ES}_\alpha (w) = \sqrt{\sum_{k=1}^5 (\mathrm{ES}_\alpha^\star (w; h_k))^2}$

Capital requirement for modellable risk factors

The final step for computing the capital requirement (also known as the 'internally modelled capital charge') is to apply this formula:

$$\mathrm{IMCC} = \varrho \cdot \mathrm{IMCC}_{global} + (1 - \varrho) \cdot \sum_{k=1}^{5} \mathrm{IMCC}_k$$

where:

- ϱ is equal to 50%
- IMCC_{global} is the stressed ES calculated with the internal model and cross-correlations between risk classes
- IMCC_k is the stressed ES calculated at the risk class level (interest rate, equity, foreign exchange, commodity and credit spread)

Other capital requirements

- Concerning non-modellable risk factors, the capital requirement is based on stress scenarios, that are equivalent to a stressed expected shortfall SES
- The default risk capital (DRC) is calculated using a value-at-risk model with a 99.9% confidence level with the same default probabilities that are used for the IRB approach

Capital requirement for the market risk

For eligible trading desks, we have:

$$\mathcal{K}_t^{\mathrm{IMA}} = \max\left(\mathrm{IMCC}_{t-1} + \mathrm{SES}_{t-1}, \frac{m_c \sum_{i=1}^{60} \mathrm{IMCC}_{t-i} + \sum_{i=1}^{60} \mathrm{SES}_{t-i}}{60}\right) + \mathrm{DRC}$$

where $m_c = 1.5 + \xi$ and $0 \leq \xi \leq 0.5$

Table: Value of the penalty coefficient ξ in Basel III

Zone	Number of exceptions	ξ
Green	0 – 4	0.00
Amber	5	0.20
	6	0.26
	7	0.33
	8	0.38
	9	0.42
Red	10+	0.50

Coherent risk measures

We note $\mathcal{R}(w)$ as the risk measure of portfolio w

Coherent risk measure

1. Subadditivity
$$\mathcal{R}(w_1 + w_2) \leq \mathcal{R}(w_1) + \mathcal{R}(w_2)$$
2. Homogeneity
$$\mathcal{R}(\lambda w) = \lambda \mathcal{R}(w) \quad \text{if } \lambda \geq 0$$
3. Monotonicity
$$\text{if } w_1 \prec w_2, \text{ then } \mathcal{R}(w_1) \geq \mathcal{R}(w_2)$$
4. Translation invariance
$$\text{if } m \in \mathbb{R}, \text{ then } \mathcal{R}(w + m) = \mathcal{R}(w) - m$$

$\Rightarrow$ Translation invariance implies that:
$$\mathcal{R}(w + \mathcal{R}(w)) = \mathcal{R}(w) - \mathcal{R}(w) = 0$$

Some risk measures

The portfolio's loss is equal to $L(w) = -P_t(w) R_{t+h}(w)$

- Volatility of the loss

$$\mathcal{R}(w) = \sigma(L(w)) = \sigma(w)$$

- Standard deviation-based risk measure

$$\mathcal{R}(w) = \mathrm{SD}_c(w) = \mathbb{E}[L(w)] + c \cdot \sigma(L(w)) = -\mu(w) + c \cdot \sigma(w)$$

- Value-at-risk

$$\mathcal{R}(w) = \mathrm{VaR}_\alpha(w) = \inf\{\ell : \Pr\{L(w) \leq \ell\} \geq \alpha\}$$

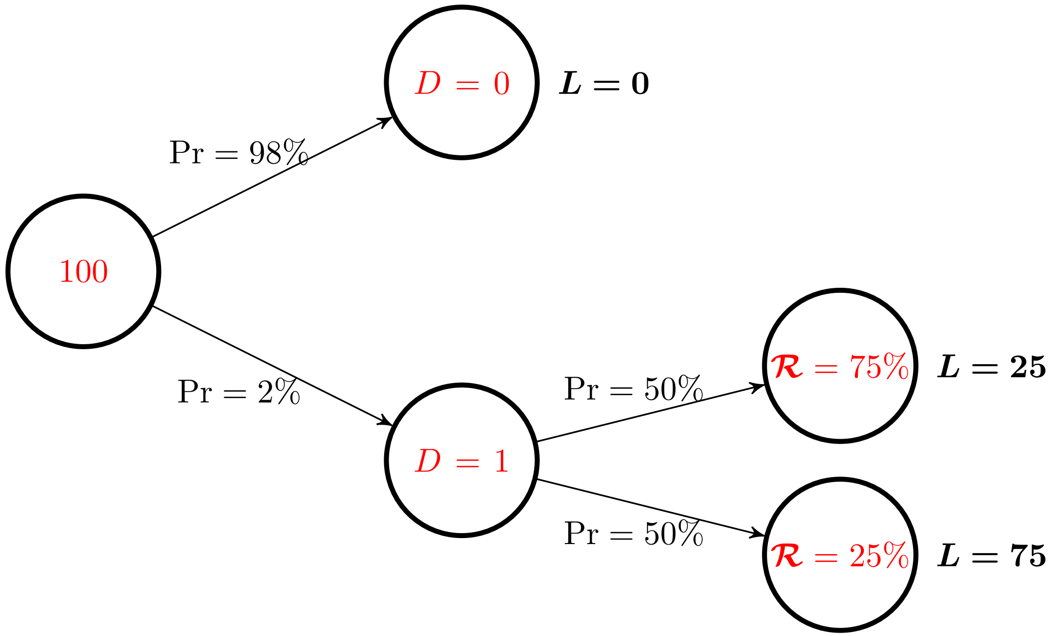
- Expected shortfall

$$\mathcal{R}(w) = \mathrm{ES}_\alpha(w) = \mathbb{E}[L(w) \mid L(w) \geq \mathrm{VaR}_\alpha(w)] = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(w)\,\mathrm{d}u$$

The value-at-risk is not always subadditive

Example

We consider a \$100 defaultable zero-coupon bond, whose default probability is equal to 200 bps. We assume that the recovery rate $\boldsymbol{\mathcal{R}}$ is a binary random variable with $\Pr\{\boldsymbol{\mathcal{R}} = 0.25\} = \Pr\{\boldsymbol{\mathcal{R}} = 0.75\} = 50\%$.



$\Rightarrow$ $\mathbf{F}(0) = \Pr\{L \leq 0\} = 98\%$, $\mathbf{F}(25) = \Pr\{L_i \leq 25\} = 99\%$ and $\mathbf{F}(75) = \Pr\{L_i \leq 75\} = 100\%$

The value-at-risk is not always subadditive

The 99% value-at-risk is equal to \$25, and we have:

$$\mathrm{ES}_{99\%}(L) = \mathbb{E}[L \mid L \geq 25] = \frac{25+75}{2} = \$50$$

We now consider two zero-coupon bonds with iid default times:

	$L_1 = 0$	$L_1 = 25$	$L_1 = 75$	
$L_2 = 0$	96.04%	0.98%	0.98%	98.00%
$L_2 = 25$	0.98%	0.01%	0.01%	1.00%
$L_2 = 75$	0.98%	0.01%	0.01%	1.00%
	98.00%	1.00%	1.00%	

We deduce that the probability distribution function of $L = L_1 + L_2$ is:

ℓ	0	25	50	75	100	150
$\Pr\{L = \ell\}$	96.04%	1.96%	0.01%	1.96%	0.02%	0.01%
$\Pr\{L \leq \ell\}$	96.04%	98%	98.01%	99.97%	99.99%	100%

It follows that $\mathrm{VaR}_{99\%}(L) = 75$ and:

$$\mathrm{ES}_{99\%}(L) = \frac{75 \times 1.96\% + 100 \times 0.02\% + 150 * 0.01\%}{1.96\% + 0.02\% + 0.01\%} = \$75.63$$

Value-at-risk

Definition

The value-at-risk $\mathrm{VaR}_\alpha(w;h)$ is defined as the potential loss which the portfolio w can suffer for a given confidence level α and a fixed holding period h:

$$\Pr\{L(w) \leq \mathrm{VaR}_\alpha(w;h)\} = \alpha \Leftrightarrow \mathrm{VaR}_\alpha(w;h) = \mathbf{F}_L^{-1}(\alpha)$$

Three parameters are necessary to compute this risk measure:

- the holding period h, which indicates the time period to calculate the loss;
- the confidence level α, which gives the probability that the loss is lower than the value-at-risk;
- the portfolio w, which gives the allocation in terms of risky assets and is related to the risk factors.

Expected shortfall

Definition

The expected shortfall $\mathrm{ES}_\alpha(w;h)$ is defined as the expected loss beyond the value-at-risk of the portfolio:

$$\mathrm{ES}_\alpha(w;h) = \mathbb{E}\left[L(w) \mid L(w) \geq \mathrm{VaR}_\alpha(w;h)\right]$$

We notice that $\mathrm{ES}_\alpha(w;h) \geq \mathrm{VaR}_\alpha(w;h)$

Three methods

Let $(\mathcal{F}_1, \ldots, \mathcal{F}_m)$ be the vector of risk factors. We assume that there is a pricing function g such that:

$$P_t(w) = g(\mathcal{F}_{1,t}, \ldots, \mathcal{F}_{m,t}; w)$$

We deduce that the expression of the random loss is equal to:

$$L(w) = P_t(w) - g(\mathcal{F}_{1,t+h}, \ldots, \mathcal{F}_{m,t+h}; w) = \ell(\mathcal{F}_{1,t+h}, \ldots, \mathcal{F}_{m,t+h}; w)$$

where ℓ is the loss function. We have:

$$\widehat{\mathrm{VaR}}_\alpha(w; h) = \hat{\mathbf{F}}_L^{-1}(\alpha) = -\hat{\mathbf{F}}_\Pi^{-1}(1-\alpha)$$

and:

$$\widehat{\mathrm{ES}}_\alpha(w; h) = \frac{1}{1-\alpha}\int_\alpha^1 \hat{\mathbf{F}}_L^{-1}(u)\,\mathrm{d}u$$

1. the historical (or empirical or non-parametric) VaR/ES
2. the analytical (or parametric or Gaussian) VaR/ES
3. the Monte Carlo (or simulated) VaR/ES

Historical methods

Two approaches:

- order statistic approach
- kernel approach

Let $(\mathcal{F}_{1,s}, \ldots, \mathcal{F}_{m,s})$ be the vector of risk factors observed at time $s < t$. If we calculate the future P&L with this historical scenario, we obtain:

$$\Pi_s(w) = g(\mathcal{F}_{1,s}, \ldots, \mathcal{F}_{m,s}; w) - P_t(w)$$

If we consider n_S historical scenarios ($s = 1, \ldots, n_S$), the empirical distribution $\hat{\mathbf{F}}_\Pi$ is described by the following probability distribution:

$\Pi(w)$	$\Pi_1(w)$	$\Pi_2(w)$	$\cdots$	$\Pi_{n_S}(w)$
p_s	$1/n_S$	$1/n_S$		$1/n_S$

Order statistic approach

Theorem (HFRM, page 67)

Let $X_1, \ldots, X_n$ be a sample from a continuous distribution $\mathbf{F}$. Suppose that for a given scalar $\alpha \in \left]0,1\right[$, there exists a sequence $\{a_n\}$ such that $\sqrt{n}\left(a_n - n\alpha\right) \rightarrow 0$. We can show that:

$$\sqrt{n}\left(X_{(a_n:n)} - \mathbf{F}^{-1}(\alpha)\right) \rightarrow \mathcal{N}\left(0, \frac{\alpha(1-\alpha)}{f^2\left(\mathbf{F}^{-1}(\alpha)\right)}\right)$$

$\Rightarrow \hat{\mathbf{F}}^{-1}(\alpha) = X_{(n\alpha:n)}$

- If $n_s = 1\,000$, $\hat{\mathbf{F}}^{-1}(90\%)$ is the 900^{th} order statistic
- If $n_s = 2\,00$, $\hat{\mathbf{F}}^{-1}(90.5\%)$ is the 181^{th} order statistic

Order statistic approach

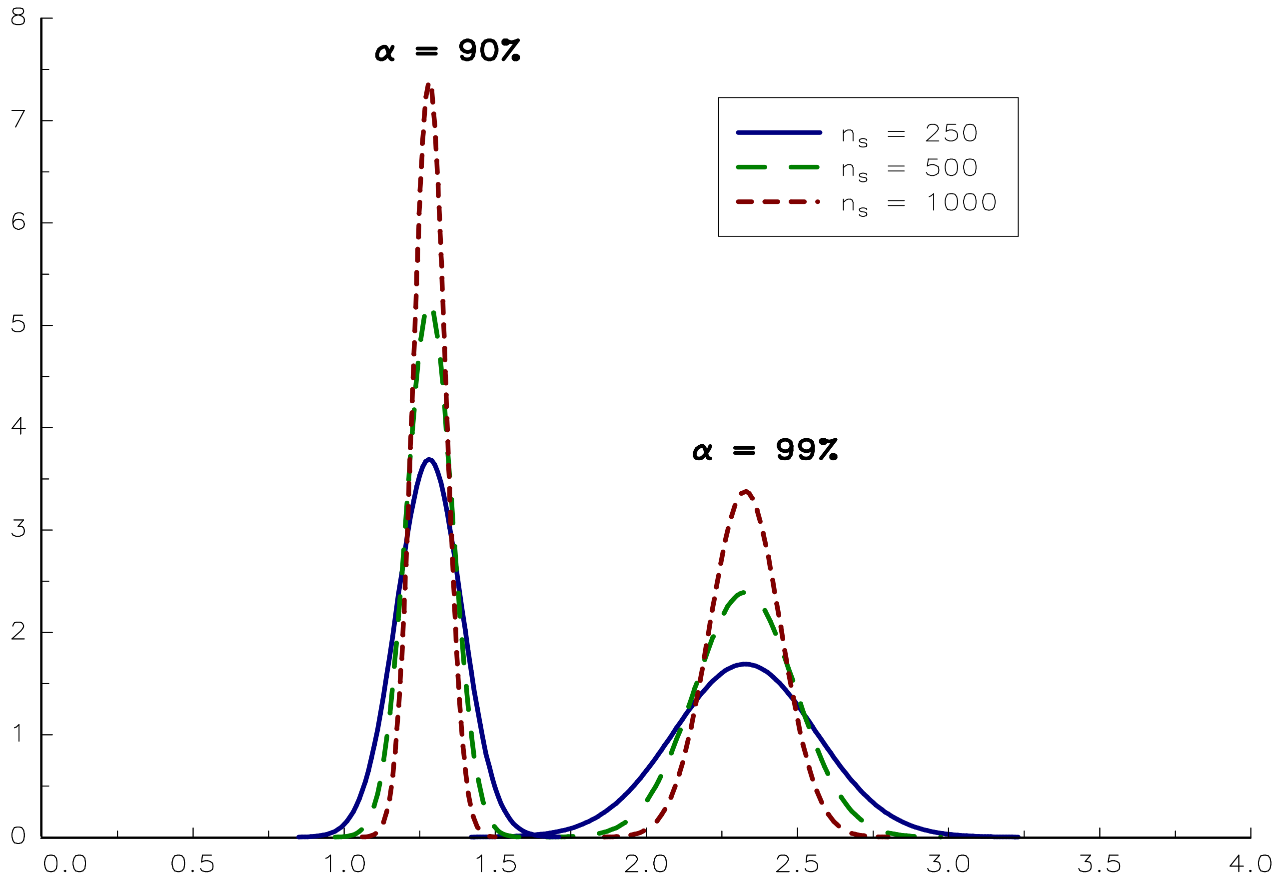


Figure: Density of the quantile estimator (Gaussian case)

Application to the value-at-risk

We calculate the order statistics associated to the P&L sample $\{\Pi_1(w), \ldots, \Pi_{n_S}(w)\}$:

$$\min_s \Pi_s(w) = \Pi_{(1:n_S)} \leq \Pi_{(2:n_S)} \leq \cdots \leq \Pi_{(n_S-1:n_S)} \leq \Pi_{(n_S:n_S)} = \max_s \Pi_s(w)$$

It follows that:

$$\mathrm{VaR}_\alpha(w; h) = -\Pi_{(n_S(1-\alpha):n_S)}$$

Application to the value-at-risk

Remark

If $n_S(1-\alpha)$ is not an integer, we consider the interpolation scheme:

$$\mathrm{VaR}_\alpha(w;h) = -\left(\Pi_{(q:n_S)} + (n_S(1-\alpha)-q)\left(\Pi_{(q+1:n_S)} - \Pi_{(q:n_S)}\right)\right)$$

where $q = q_\alpha(n_S) = \lfloor n_S(1-\alpha)\rfloor$ is the integer part of $n_S(1-\alpha)$.

In the case where we use 250 historical scenarios, the 99% value-at-risk is the mean between the second and third largest losses:

$$\begin{aligned}
\mathrm{VaR}_{99\%}(w;h) &= -\left(\Pi_{(2:250)} + (2.5-2)\left(\Pi_{(3:250)} - \Pi_{(2:250)}\right)\right) \\
&= -\frac{1}{2}\left(\Pi_{(2:250)} + \Pi_{(3:250)}\right) \\
&= \frac{1}{2}\left(L_{(249:250)} + L_{(248:250)}\right)
\end{aligned}$$

Application to the value-at-risk

Example

We consider a portfolio composed of 10 stocks Apple and 20 stocks Coca-Cola. The current date is 2 January 2015.

Remark

Data are available at
http://www.thierry-roncalli.com/download/frm-data1.xlsx

Application to the value-at-risk

The mark-to-market of the portfolio is:

$$P_t(w) = 10 \times P_{1,t} + 20 \times P_{2,t}$$

where $P_{1,t}$ and $P_{2,t}$ are the stock prices of Apple and Coca-Cola. We assume that the market risk factors corresponds to the daily stock returns $R_{1,t}$ and $R_{2,t}$. We deduce that the P&L for the scenario s is equal to:

$$\Pi_s(w) = \underbrace{10 \times P_{1,s} + 20 \times P_{2,s}}_{g(R_{1,s}, R_{2,s}; w)} - P_t(w)$$

where $P_{i,s} = P_{i,t} \times (1 + R_{i,s})$ is the simulated price of stock i for the scenario s.

Application to the value-at-risk

Table: Computation of the market risk factors $R_{1,s}$ and $R_{2,s}$

s	Date	Apple		Coca-Cola	
		Price	$R_{1,s}$	Price	$R_{2,s}$
1	2015-01-02	109.33	−0.95%	42.14	−0.19%
2	2014-12-31	110.38	−1.90%	42.22	−1.26%
3	2014-12-30	112.52	−1.22%	42.76	−0.23%
4	2014-12-29	113.91	−0.07%	42.86	−0.23%
5	2014-12-26	113.99	1.77%	42.96	0.05%
6	2014-12-24	112.01	−0.47%	42.94	−0.07%
7	2014-12-23	112.54	−0.35%	42.97	1.46%
8	2014-12-22	112.94	1.04%	42.35	0.95%
9	2014-12-19	111.78	−0.77%	41.95	−1.04%
10	2014-12-18	112.65	2.96%	42.39	2.02%

Application to the value-at-risk

- We calculate the historical risk factors. For instance, we have:

$$R_{1,1} = \frac{109.33}{110.38} - 1 = -0.95\%$$

- We calculate the simulated prices. For instance, in the case of the 9^{th} scenario, we obtain:

$$\begin{aligned} P_{1,s} &= 109.33 \times (1 - 0.77\%) = \$108.49 \\ P_{2,s} &= 42.14 \times (1 - 1.04\%) = \$41.70 \end{aligned}$$

- We then deduce the simulated mark-to-market $\mathrm{MtM}_s(w) = g(R_{1,s}, R_{2,s}; w)$

Application to the value-at-risk

Table: Computation of the simulated P&L $\Pi_s(w)$

s	Date	Apple $R_{1,s}$	Apple $P_{1,s}$	Coca-Cola $R_{2,s}$	Coca-Cola $P_{2,s}$	$\mathrm{MtM}_s(w)$	$\Pi_s(w)$
1	2015-01-02	−0.95%	108.29	−0.19%	42.06	1 924.10	−12.00
2	2014-12-31	−1.90%	107.25	−1.26%	41.61	1 904.66	−31.44
3	2014-12-30	−1.22%	108.00	−0.23%	42.04	1 920.79	−15.31
4	2014-12-29	−0.07%	109.25	−0.23%	42.04	1 933.37	−2.73
5	2014-12-26	1.77%	111.26	0.05%	42.16	1 955.82	19.72
23	2014-12-01	−3.25%	105.78	−0.62%	41.88	1 895.35	−40.75
69	2014-09-25	−3.81%	105.16	−1.16%	41.65	1 884.64	−51.46
85	2014-09-03	−4.22%	104.72	0.34%	42.28	1 892.79	−43.31
108	2014-07-31	−2.60%	106.49	−0.83%	41.79	1 900.68	−35.42
236	2014-01-28	−7.99%	100.59	0.36%	42.29	1 851.76	−84.34
242	2014-01-17	−2.45%	106.65	−1.08%	41.68	1 900.19	−35.91
250	2014-01-07	−0.72%	108.55	0.30%	42.27	1 930.79	−5.31

Application to the value-at-risk

If we rank the scenarios, the worst P&Ls are -84.34, -51.46, -43.31, -40.75, -35.91 and -35.42. We deduce that the daily historical VaR is equal to:

$$\mathrm{VaR}_{99\%}(w; \text{one day}) = \frac{1}{2}(51.46 + 43.31) = \$47.39$$

If we assume that $m_c = 3$, the corresponding capital charge represents 23.22% of the portfolio value:

$$\mathcal{K}_t^{\mathrm{VaR}} = 3 \times \sqrt{10} \times 47.39 = \$449.54$$

Application to the expected shortfall

Since the expected shortfall is the expected loss beyond the value-at-risk, it follows that the historical expected shortfall is given by:

$$\mathrm{ES}_\alpha(w;h) = \frac{1}{q_\alpha(n_S)} \sum_{s=1}^{n_S} \mathbb{1}\{L_s \geq \mathrm{VaR}_\alpha(w;h)\} \cdot L_s$$

or:

$$\mathrm{ES}_\alpha(w;h) = -\frac{1}{q_\alpha(n_S)} \sum_{s=1}^{n_S} \mathbb{1}\{\Pi_s \leq -\mathrm{VaR}_\alpha(w;h)\} \cdot \Pi_s$$

where $q_\alpha(n_S) = \lfloor n_s(1-\alpha) \rfloor$ is the integer part of $n_s(1-\alpha)$.

Computation of the ES

We have:

$$\mathrm{ES}_\alpha(w;h) = -\frac{1}{q_\alpha(n_S)} \sum_{i=1}^{q_\alpha(n_S)} \Pi_{(i:n_S)}$$

Application to the expected shortfall

We have:

$$\mathrm{ES}_{99\%}(w; \text{one day}) = \frac{84.34 + 51.46}{2} = \$67.90$$

and:

$$\mathrm{ES}_{97.5\%}(w; \text{one day}) = \frac{84.34 + 51.46 + 43.31 + 40.75 + 35.91 + 35.42}{6} = \$48.53$$

We remind that $\mathrm{VaR}_{99\%}(w; \text{one day}) = \$47.39$.

Analytical methods

We speak about analytical value-at-risk when we are able to find a closed-form formula of $\mathbf{F}_L^{-1}(\alpha)$

Gaussian value-at-risk

Suppose that $L(w) \sim \mathcal{N}\left(\mu(L), \sigma^2(L)\right)$. In this case, we have $\Pr\left\{L(w) \leq \mathbf{F}_L^{-1}(\alpha)\right\} = \alpha$ or:

$$\Pr\left\{\frac{L(w) - \mu(L)}{\sigma(L)} \leq \frac{\mathbf{F}_L^{-1}(\alpha) - \mu(L)}{\sigma(L)}\right\} = \alpha \Leftrightarrow \Phi\left(\frac{\mathbf{F}_L^{-1}(\alpha) - \mu(L)}{\sigma(L)}\right) = \alpha$$

We deduce that:

$$\frac{\mathbf{F}_L^{-1}(\alpha) - \mu(L)}{\sigma(L)} = \Phi^{-1}(\alpha) \Leftrightarrow \mathbf{F}_L^{-1}(\alpha) = \mu(L) + \Phi^{-1}(\alpha)\,\sigma(L)$$

The expression of the value-at-risk is then:

$$\mathrm{VaR}_\alpha(w; h) = \mu(L) + \Phi^{-1}(\alpha)\,\sigma(L)$$

if $\alpha = 99\%$, we obtain:

$$\mathrm{VaR}_{99\%}(w; h) = \mu(L) + 2.33 \times \sigma(L)$$

Gaussian value-at-risk

Example

We consider a short position of \$1 mn on the S&P 500 futures contract. We estimate that the annualized volatility $\hat{\sigma}_{\mathrm{SPX}}$ is equal to 35%

The portfolio loss is equal to $L(w) = N \times R_{\mathrm{SPX}}$ where N is the exposure amount ($-\$1$ mn) and R_{SPX} is the (Gaussian) return of the S&P 500 index. We deduce that the annualized loss volatility is $\hat{\sigma}(L) = |N| \times \hat{\sigma}_{\mathrm{SPX}}$. The value-at-risk for a one-year holding period is:

$$\mathrm{VaR}_{99\%}(w; \text{one year}) = 2.33 \times 10^6 \times 0.35 = \$815\,500$$

By using the square-root-of-time rule, we deduce that:

$$\mathrm{VaR}_{99\%}(w; \text{one day}) = \frac{815\,500}{\sqrt{260}} = \$50\,575$$

Gaussian expected shortfall

By definition, we have:

$$\begin{aligned}\mathrm{ES}_\alpha(w) &= \mathbb{E}\left[L(w) \mid L(w) \geq \mathrm{VaR}_\alpha(w)\right] \\ &= \frac{1}{1-\alpha}\int_{\mathbf{F}_L^{-1}(\alpha)}^{\infty} x f_L(x)\,\mathrm{d}x\end{aligned}$$

where f_L and $\mathbf{F}_L$ are the density and distribution functions of the loss $L(w)$

The Gaussian expected shortfall of the portfolio w is equal to:

$$\mathrm{ES}_\alpha(w) = \mu(L) + \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{(1-\alpha)}\sigma(L)$$

Proof

$$\mathrm{ES}_{\alpha}(w) = \frac{1}{1-\alpha} \int_{\mu(L)+\Phi^{-1}(\alpha)\sigma(L)}^{\infty} \frac{x}{\sigma(L)\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu(L)}{\sigma(L)}\right)^2\right) \mathrm{d}x$$

With the variable change $t = \sigma(L)^{-1}(x - \mu(L))$, we obtain:

$$\begin{aligned}
\mathrm{ES}_{\alpha}(w) &= \frac{1}{1-\alpha} \int_{\Phi^{-1}(\alpha)}^{\infty} (\mu(L) + \sigma(L)\, t) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) \mathrm{d}t \\
&= \frac{\mu(L)}{1-\alpha} \left[\Phi(t)\right]_{\Phi^{-1}(\alpha)}^{\infty} + \frac{\sigma(L)}{(1-\alpha)\sqrt{2\pi}} \int_{\Phi^{-1}(\alpha)}^{\infty} t \exp\left(-\frac{1}{2}t^2\right) \mathrm{d}t \\
&= \mu(L) + \frac{\sigma(L)}{(1-\alpha)\sqrt{2\pi}} \left[-\exp\left(-\frac{1}{2}t^2\right)\right]_{\Phi^{-1}(\alpha)}^{\infty} \\
&= \mu(L) + \frac{\sigma(L)}{(1-\alpha)\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\Phi^{-1}(\alpha)\right]^2\right)
\end{aligned}$$

Gaussian VaR vs Gaussian ES

The value-at-risk and the expected shortfall are both a standard deviation-based risk measure. They coincide when the scaling parameters $c_{\mathrm{VaR}} = \Phi^{-1}(\alpha_{\mathrm{VaR}})$ and $c_{\mathrm{ES}} = \phi\left(\Phi^{-1}(\alpha_{\mathrm{ES}})\right) / (1 - \alpha_{\mathrm{ES}})$ are equal.

Table: Scaling factors c_{VaR} and c_{ES}

α (in %)	95.0	96.0	97.0	97.5	98.0	98.5	99.0	99.5
c_{VaR}	1.64	1.75	1.88	1.96	2.05	2.17	**2.33**	2.58
c_{ES}	2.06	2.15	2.27	**2.34**	2.42	2.52	2.67	2.89

Linear factor models

When $g(\mathcal{F}_t; w) = \sum_{i=1}^{n} w_i P_{i,t}$, the random P&L is equal to:

$$\begin{aligned}\Pi(w) &= P_{t+h}(w) - P_t(w) \\ &= \sum_{i=1}^{n} w_i P_{i,t+h} - \sum_{i=1}^{n} w_i P_{i,t} \\ &= \sum_{i=1}^{n} w_i \left(P_{i,t+h} - P_{i,t}\right)\end{aligned}$$

We assume that the asset returns are the risk factors :

$$P_{i,t+h} = P_{i,t}\left(1 + R_{i,t+h}\right)$$

where $R_{i,t+h}$ is the asset return between t and $t+h$. In this case, we obtain:

$$\Pi(w) = \sum_{i=1}^{n} w_i P_{i,t} R_{i,t+h}$$

The covariance model

Let R_t be the vector of asset returns. We note $W_{i,t} = w_i P_{i,t}$ the wealth invested (or the nominal exposure) in asset i and $W_t = (W_{1,t}, \ldots, W_{n,t})$. It follows that:

$$\Pi(w) = \sum_{i=1}^{n} W_{i,t} R_{i,t+h} = W_t^\top R_{t+h}$$

If we assume that $R_{t+h} \sim \mathcal{N}(\mu, \Sigma)$, we deduce that $\mu(\Pi) = W_t^\top \mu$ and $\sigma^2(\Pi) = W_t^\top \Sigma W_t$. Therefore, the expression of the value-at-risk is:

$$\mathrm{VaR}_\alpha(w; h) = -W_t^\top \mu + \Phi^{-1}(\alpha)\sqrt{W_t^\top \Sigma W_t}$$

Example

We consider the Apple/Coca-Cola example. The nominal exposures are \$1 093.3 (Apple) and \$842.8 (Coca-Cola). The estimated standard deviation of daily returns is equal to 1.3611% for Apple and 0.9468% for Coca-Cola, whereas the cross-correlation is equal to 12.0787%. It follows that:

$$\begin{aligned}\sigma^2(\Pi) &= W_t^\top \Sigma W_t \\ &= 1\,093.3^2 \times \left(\frac{1.3611}{100}\right)^2 + 842.8^2 \times \left(\frac{0.9468}{100}\right)^2 + \\ &\quad 2 \times \frac{12.0787}{100} \times 1\,093.3 \times 842.8 \times \frac{1.3611}{100} \times \frac{0.9468}{100} \\ &= 313.80\end{aligned}$$

We deduce that the 99% daily value-at-risk is equal to:

$$\mathrm{VaR}_{99\%}(w; \text{one day}) = \Phi^{-1}(0.99)\sqrt{313.80} = \$41.21$$

The factor model

- CAPM (HFRM, pages 76-77)
- APT (HFRM, page 77 and Exercise 2.4.5 page 119)
- Application to a bond portfolio (HFRM, pages 77-80)

Some other topics

- Volatility forecasting EWMA, GARCH and SV models (HFRM, pages 80-83 and Section 10.2.4 page 664)
- Other probability distributions (HFRM, pages 84-90)
- Cornish-Fisher approximation (HFRM, pages 85-87)

$$\mathrm{VaR}_{\alpha}(w;h) = \mu(L) + Z(\alpha;\gamma_1(L),\gamma_2(L)) \times \sigma(L)$$

where:

$$Z(\alpha;\gamma_1,\gamma_2) = z_\alpha + \frac{1}{6}\left(z_\alpha^2 - 1\right)\gamma_1 + \frac{1}{24}\left(z_\alpha^3 - 3z_\alpha\right)\gamma_2 - \frac{1}{36}\left(2z_\alpha^3 - 5z_\alpha\right)\gamma_1^2$$

and $z_\alpha = \Phi^{-1}(\alpha)$

Monte Carlo methods

- We assume a given probability distribution $\mathbf{H}$ for the risk factors:

$$\left(\mathcal{F}_{1,t+h}, \ldots, \mathcal{F}_{m,t+h}\right) \sim \mathbf{H}$$

- We simulate n_S scenarios of risk factors and calculate the simulated P&L $\Pi_s(w)$ for each scenario s
- We calculate the empirical quantile using the order statistic approach

$\Rightarrow$ The **Monte Carlo VaR/ES** is a **historical VaR/ES** with simulated scenarios or the Monte Carlo VaR/ES is a **parametric VaR/ES** for which it is difficult to find an analytical formula

Identification of risk factors

We consider a portfolio containing w_S stocks and w_C call options on this stock. We note S_t and $\mathcal{C}_t$ the stock and option prices at time t. We have:

$$\Pi(w) = w_S(S_{t+h} - S_t) + w_C(\mathcal{C}_{t+h} - \mathcal{C}_t)$$

If we use asset returns as risk factors, we get:

$$\Pi(w) = w_S S_t R_{S,t+h} + w_C \mathcal{C}_t R_{C,t+h}$$

where $R_{S,t+h}$ and $R_{C,t+h}$ are the returns of the stock and the option for the period $[t, t+h]$

$\Rightarrow$ **Two risk factors**: $R_{S,t+h}$ and $R_{C,t+h}$?

Identification of risk factors

The problem is that the option price $\mathcal{C}_t$ is a non-linear function of the underlying price S_t:

$$\mathcal{C}_t = f_C(S_t)$$

This implies that:

$$\begin{aligned}\Pi(w) &= w_S S_t R_{S,t+h} + w_C \left(f_C(S_{t+h}) - \mathcal{C}_t\right) \\ &= w_S S_t R_{S,t+h} + w_C \left(f_C(S_t(1 + R_{S,t+h})) - \mathcal{C}_t\right)\end{aligned}$$

$\Rightarrow$ **One risk factor**: $R_{S,t+h}$?

The Black-Scholes formula

The price of the call option is equal to:

$$C_{\mathrm{BS}}\left(S_t, K, \Sigma_t, T, b_t, r_t\right) = S_t e^{(b_t - r_t)\tau} \Phi\left(d_1\right) - K e^{-r_t \tau} \Phi\left(d_2\right)$$

where:

- S_t is the current price of the underlying asset
- K is the option strike
- Σ_t is the volatility parameter,
- T is the maturity date
- b_t is the cost-of-carry[9]
- r_t is the interest rate
- the parameter $\tau = T - t$ is the time to maturity
- The coefficients d_1 and d_2 are defined as follows:

$$d_1 = \frac{1}{\Sigma_t \sqrt{\tau}} \left(\ln \frac{S_t}{K} + b_t \tau \right) + \frac{1}{2} \Sigma_t \sqrt{\tau} \quad \text{and} \quad d_2 = d_1 - \Sigma_t \sqrt{\tau}$$

[9] The cost-of-carry depends on the underlying asset. We have $b_t = r_t$ for non-dividend stocks and total return indices, $b_t = r_t - d_t$ for stocks paying a continuous dividend yield d_t, $b_t = 0$ for forward and futures contracts and $b_t = r_t - r_t^{\star}$ for foreign exchange options where $r_t^{\star}$ is the foreign interest rate.

Identification of risk factors

We can write the option price as follows:

$$\mathcal{C}_t = f_{\mathrm{BS}}\left(\theta_{\mathrm{contract}}; \theta\right)$$

where $\theta_{\mathrm{contract}}$ are the parameters of the contract (strike K and maturity T) and θ are the other parameters

- S_t is obviously a risk factor
- If Σ_t is not constant, the option price may be sensitive to the volatility risk
- The option may be impacted by changes in the interest rate or the cost-of-carry

$\Rightarrow$ The choice of risk factors depends on the derivative contract (volatility risk, dividend risk, yield curve risk, correlation risk, etc.)

Methods to calculate VAR and ES risk measures

1. The method of full pricing (option repricing)
2. The method of sensitivities (delta-gamma-vega approximation)
3. The hybrid method

The method of full pricing

We recall that the P&L of the s^{th} scenario has the following expression:

$$\Pi_s(w) = g(\mathcal{F}_{1,s}, \ldots, \mathcal{F}_{m,s}; w) - P_t(w)$$

In the case of the previous example, the P&L becomes then:

$$\Pi_s(w) = \begin{cases} w_S S_t R_s + w_C \left(f_C\left(S_t(1+R_s); \Sigma_t\right) - \mathcal{C}_t\right) & \text{with one risk factor} \\ w_S S_t R_s + w_C \left(f_C\left(S_t(1+R_s), \Sigma_s\right) - \mathcal{C}_t\right) & \text{with two risk factors} \end{cases}$$

where the pricing function is:

$$f_C(S; \Sigma) = C_{\mathrm{BS}}(S, K, \Sigma, T - h, b_t, r_t)$$

Remark

In the model with two risk factors, we have to simulate the underlying price and the implied volatility. For the single factor model, we use the current implied volatility Σ_t instead of the simulated value Σ_s.

Application to the VaR and ES

Example

We consider a long position on 100 call options with strike $K = 100$. The value of the call option is \$4.14, the residual maturity is 52 days and the current price of the underlying asset is \$100. We assume that $\Sigma_t = 20\%$ and $b_t = r_t = 5\%$. The objective is to calculate the daily 99% VaR and the daily 97.5% ES with 250 historical scenarios, whose first nine values are the following:

s	1	2	3	4	5	6	7	8	9
R_s	−1.93	−0.69	−0.71	−0.73	1.22	1.01	1.04	1.08	−1.61
$\Delta\Sigma_s$	−4.42	−1.32	−3.04	2.88	−0.13	−0.08	1.29	2.93	0.85

Remark

Data are available at
http://www.thierry-roncalli.com/download/frm-data1.xlsx

Application to the VaR and ES

$\Rightarrow$ The implied volatility is equal to 20%

For the first scenario, R_s is equal to -1.93% and S_{t+h} is equal to $100 \times (1 - 1.93\%) = 98.07$. The residual maturity τ is equal to $51/252$ years. It follows that:

$$d_1 = \frac{1}{20\% \times \sqrt{51/252}} \left(\ln \frac{98.07}{100} + 5\% \times \frac{51}{252} \right) + \frac{1}{2} \times 20\% \times \sqrt{\frac{51}{252}} = -0.0592$$

$$d_2 = -0.0592 - 20\% \times \sqrt{\frac{51}{252}} = -0.1491$$

We deduce that:

$$\begin{aligned} \mathcal{C}_{t+h} &= 98.07 \times e^{(5\%-5\%)\frac{51}{252}} \times \Phi(-0.0592) - 100 \times e^{5\% \times \frac{51}{252}} \times \Phi(-0.1491) \\ &= 98.07 \times 1.00 \times 0.4764 - 100 \times 1.01 \times 0.4407 \\ &= 3.093 \end{aligned}$$

The simulated P&L for the first historical scenario is then equal to:

$$\Pi_s = 100 \times (3.093 - 4.14) = -104.69$$

Application to the VaR and ES

Table: Daily P&L of the long position on the call option when the risk factor is the underlying price

s	R_s (in %)	S_{t+h}	$\mathcal{C}_{t+h}$	Π_s
1	-1.93	98.07	3.09	-104.69
2	-0.69	99.31	3.72	-42.16
3	-0.71	99.29	3.71	-43.22
4	-0.73	99.27	3.70	-44.28
5	1.22	101.22	4.81	67.46
6	1.01	101.01	4.68	54.64
7	1.04	101.04	4.70	56.46
8	1.08	101.08	4.73	58.89
9	-1.61	98.39	3.25	-89.22

$\Rightarrow$ With the 250 historical scenarios, the 99% value-at-risk is equal to \$154.79, whereas the 97.5% expected shortfall is equal to \$150.04

The option return R_C is not a new risk factor

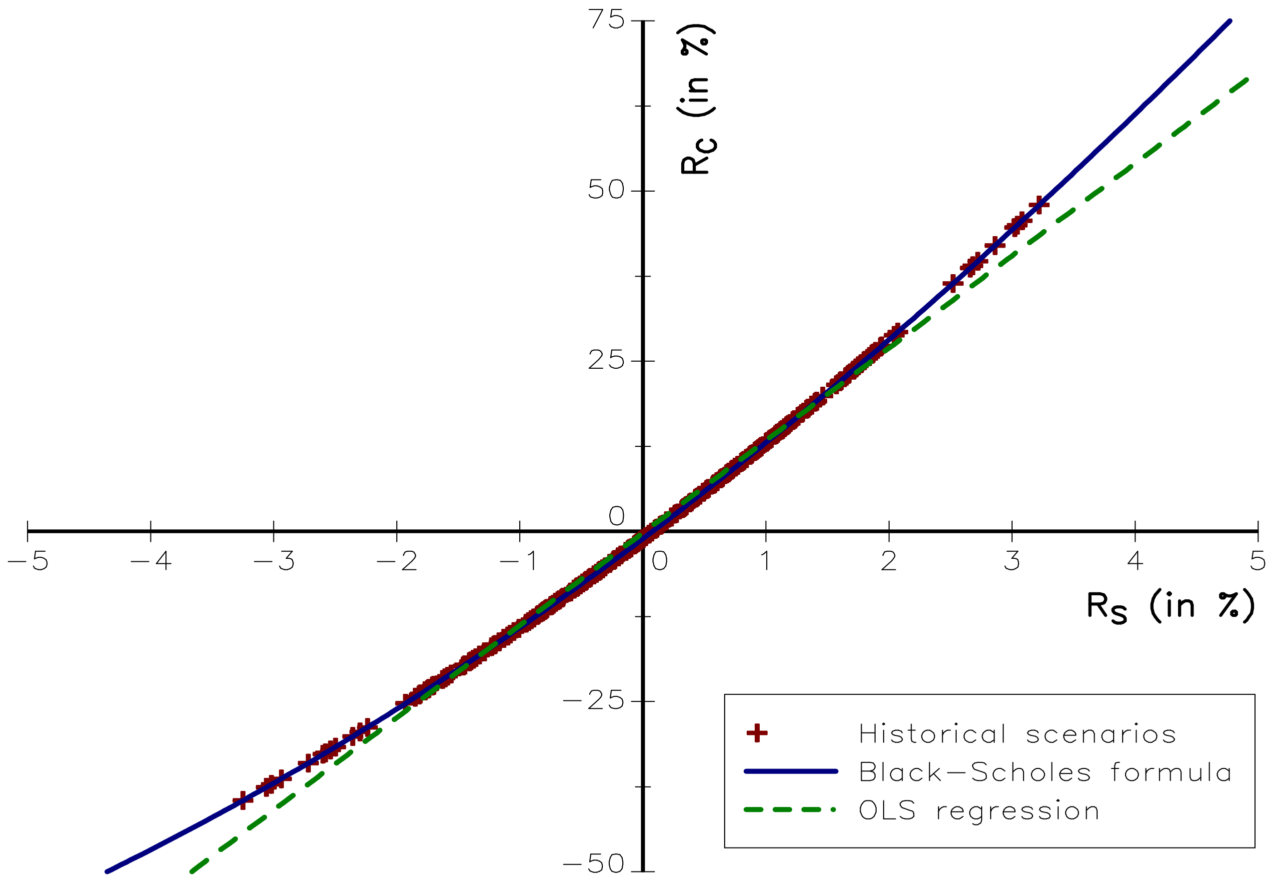


Figure: Relationship between the asset return R_S and the option return R_C

Adding the risk factor Σ_t

$$\Sigma_{t+h} = \Sigma_t + \Delta\Sigma_s$$

Table: Daily P&L of the long position on the call option when the risk factors are the underlying price and the implied volatility

s	R_s (in %)	S_{t+h}	$\Delta\Sigma_s$ (in %)	Σ_{t+h}	$\mathcal{C}_{t+h}$	Π_s
1	−1.93	98.07	−4.42	15.58	2.32	−182.25
2	−0.69	99.31	−1.32	18.68	3.48	−65.61
3	−0.71	99.29	−3.04	16.96	3.17	−97.23
4	−0.73	99.27	2.88	22.88	4.21	6.87
5	1.22	101.22	−0.13	19.87	4.79	65.20
6	1.01	101.01	−0.08	19.92	4.67	53.24
7	1.04	101.04	1.29	21.29	4.93	79.03
8	1.08	101.08	2.93	22.93	5.24	110.21
9	−1.61	98.39	0.85	20.85	3.40	−74.21

$\Rightarrow \mathrm{VaR}_{99\%}(w; \text{one day}) = \$181.70$ and $\mathrm{ES}_{97.5\%}(w; \text{one day}) = \$172.09$

The method of sensitivities

The previous approach is called *full pricing*, because it consists in re-pricing the option

In the method based on the Greek coefficients, the idea is to approximate the change in the option price by a Taylor expansion:

- Delta approach
- Delta-gamma approach
- Delta-gamma-theta approach
- Delta-gamma-theta-vega approach
- Etc.

The delta approach

We define the delta approach as follows:

$$\mathcal{C}_{t+h} - \mathcal{C}_t \simeq \boldsymbol{\Delta}_t \left(S_{t+h} - S_t \right)$$

where $\boldsymbol{\Delta}_t$ is the option delta:

$$\boldsymbol{\Delta}_t = \frac{\partial \, C_{\mathrm{BS}} \left(S_t, \Sigma_t, T \right)}{\partial \, S_t}$$

The delta approach applied to delta neutral portfolios

If we consider the introductory example, we have:

$$\begin{aligned}\Pi(w) &= w_S (S_{t+h} - S_t) + w_C (\mathcal{C}_{t+h} - \mathcal{C}_t) \\ &\simeq (w_S + w_C \mathbf{\Delta}_t)(S_{t+h} - S_t) \\ &= (w_S + w_C \mathbf{\Delta}_t) S_t R_{S,t+h}\end{aligned}$$

With the delta approach, we aggregate the risk by netting the different delta exposures[10]. In particular, the portfolio is delta neutral if the net exposure is zero:

$$w_S + w_C \mathbf{\Delta}_t = 0 \Leftrightarrow w_S = -w_C \mathbf{\Delta}_t$$

With the delta approach, the VaR/ES of delta neutral portfolios is then equal to zero

[10] A long (or short) position on the underlying asset is equivalent to $\mathbf{\Delta}_t = 1$ (or $\mathbf{\Delta}_t = -1$).

The delta-gamma approach

We can use the second-order approximation or the delta-gamma approach:

$$\mathcal{C}_{t+h} - \mathcal{C}_t \simeq \boldsymbol{\Delta}_t \left(S_{t+h} - S_t\right) + \frac{1}{2} \boldsymbol{\Gamma}_t \left(S_{t+h} - S_t\right)^2$$

where $\boldsymbol{\Gamma}_t$ is the option gamma:

$$\boldsymbol{\Gamma}_t = \frac{\partial^2 \, C_{\mathrm{BS}} \left(S_t, \Sigma_t, T\right)}{\partial \, S_t^2}$$

Comparison between delta and delta-gamma approaches

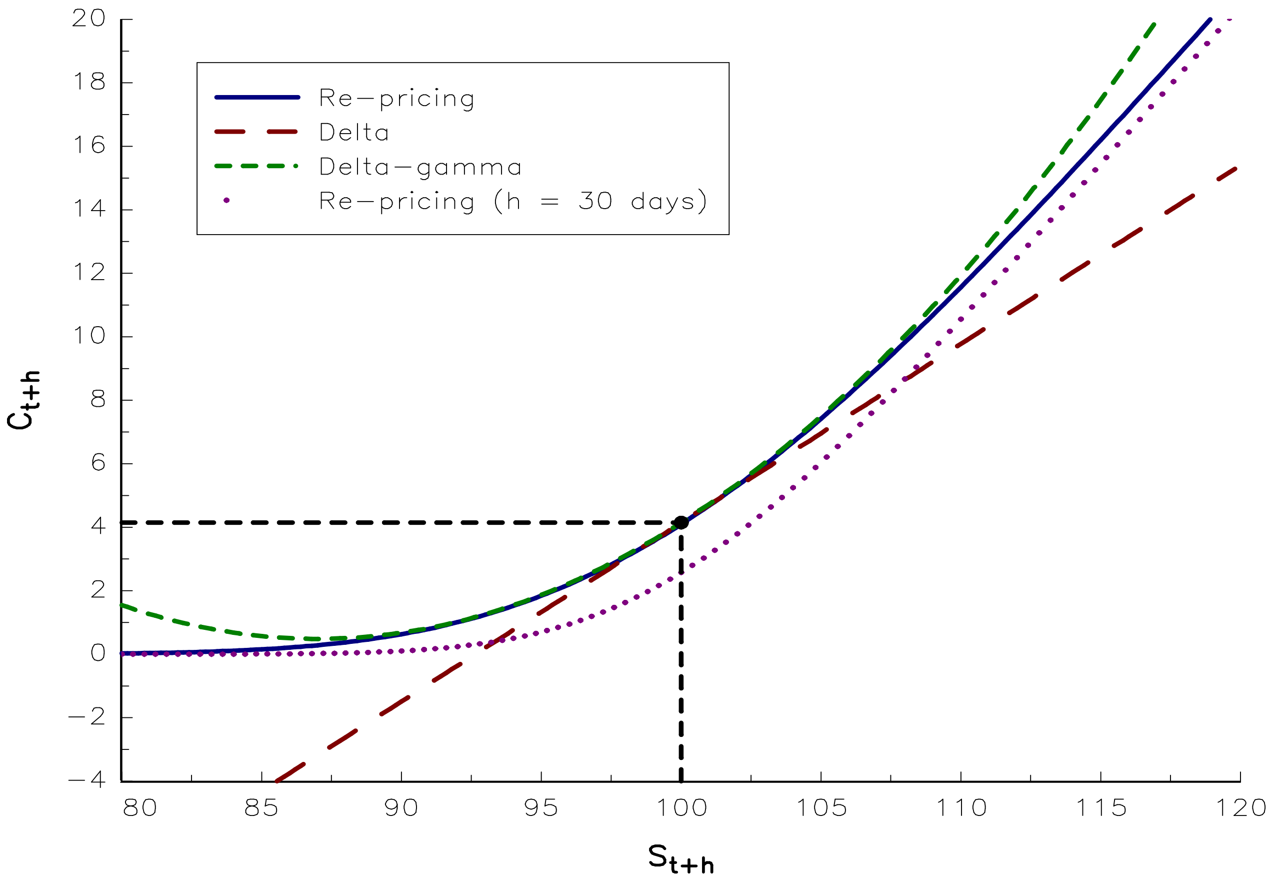


Figure: Approximation of the option price with the Greek coefficients

Extension to other risk factors

The Taylor expansion can be generalized to a set of risk factors $\mathcal{F}_t = (\mathcal{F}_{1,t}, \ldots, \mathcal{F}_{m,t})$:

$$\begin{aligned}\mathcal{C}_{t+h} - \mathcal{C}_t &\simeq \sum_{j=1}^{m} \frac{\partial\, \mathcal{C}_t}{\partial\, \mathcal{F}_{j,t}} \left(\mathcal{F}_{j,t+h} - \mathcal{F}_{j,t}\right) + \\ &\quad \frac{1}{2} \sum_{j=1}^{m} \sum_{k=1}^{m} \frac{\partial^2\, \mathcal{C}_t}{\partial\, \mathcal{F}_{j,t}\, \partial\, \mathcal{F}_{k,t}} \left(\mathcal{F}_{j,t+h} - \mathcal{F}_{j,t}\right)\left(\mathcal{F}_{k,t+h} - \mathcal{F}_{k,t}\right)\end{aligned}$$

The delta-gamma-theta-vega approach is defined as follows:

$$\mathcal{C}_{t+h} - \mathcal{C}_t \simeq \boldsymbol{\Delta}_t \left(S_{t+h} - S_t\right) + \frac{1}{2} \boldsymbol{\Gamma}_t \left(S_{t+h} - S_t\right)^2 + \boldsymbol{\Theta}_t h + \boldsymbol{\upsilon}_t \left(\Sigma_{t+h} - \Sigma_t\right)$$

where $\boldsymbol{\Theta}_t = \partial_t\, C_{\mathrm{BS}}\left(S_t, \Sigma_t, T\right)$ is the option theta and $\boldsymbol{\upsilon}_t = \partial_{\Sigma_t}\, C_{\mathrm{BS}}\left(S_t, \Sigma_t, T\right)$ is the option vega

$\Rightarrow$ We can also include vanna and volga effects

The Black-Scholes Greek coefficients

$$
\begin{aligned}
\boldsymbol{\Delta}_t &= e^{(b_t - r_t)\tau} \Phi(d_1) \\
\boldsymbol{\Gamma}_t &= \frac{e^{(b_t - r_t)\tau} \phi(d_1)}{S_t \Sigma_t \sqrt{\tau}} \\
\boldsymbol{\Theta}_t &= -r_t K e^{-r_t \tau} \Phi(d_2) - \frac{1}{2\sqrt{\tau}} S_t \Sigma_t e^{(b_t - r_t)\tau} \phi(d_1) - \\
&\quad (b_t - r_t) S_t e^{(b_t - r_t)\tau} \Phi(d_1) \\
\boldsymbol{\upsilon}_t &= e^{(b_t - r_t)\tau} S_t \sqrt{\tau} \phi(d_1)
\end{aligned}
$$

(HFRM, Exercise 2.4.7 page 121)

Application to the VaR and ES

In the case of our previous example (Slide 82), we obtain $\boldsymbol{\Delta}_t = 0.5632$, $\boldsymbol{\Gamma}_t = 0.0434$, $\boldsymbol{\Theta}_t = -11.2808$ and $\boldsymbol{\upsilon}_t = 17.8946$

We have:

- $\Pi_1^{\boldsymbol{\Delta}}(w) = 100 \times 0.5632 \times (98.07 - 100) = -108.69$
- $\Pi_1^{\boldsymbol{\Delta}+\boldsymbol{\Gamma}}(w) = -108.69 + 100 \times \frac{1}{2} \times 0.0434 \times (98.07 - 100)^2 = -100.61$
- $\Pi_1^{\boldsymbol{\Delta}+\boldsymbol{\Gamma}+\boldsymbol{\Theta}}(w) = -100.61 - 11.2808 \times 1/252 = -105.09$
- $\Pi_1^{\boldsymbol{\upsilon}}(w) = 100 \times 17.8946 \times (15.58\% - 20\%) = -79.09$
- $\Pi_1^{\boldsymbol{\Delta}+\boldsymbol{\Gamma}+\boldsymbol{\Theta}+\boldsymbol{\upsilon}}(w) = -105.90 - 79.09 = -184.99$

Application to the VaR and ES

Table: Calculation of the P&L based on the Greek sensitivities

s	R_s (in %)	S_{t+h}	Π_s	Π_s^{Δ}	$\Pi_s^{\Delta+\Gamma}$	$\Pi_s^{\Delta+\Gamma+\Theta}$
1	−1.93	98.07	−104.69	−108.69	−100.61	−105.09
2	−0.69	99.31	−42.16	−38.86	−37.83	42.30
3	−0.71	99.29	−43.22	−39.98	−38.89	−43.37
4	−0.73	99.27	−44.28	−41.11	−39.96	−44.43
5	1.22	101.22	67.46	68.71	71.93	67.46
6	1.01	101.01	54.64	56.88	59.09	54.61
7	1.04	101.04	56.46	58.57	60.91	56.44
8	1.08	101.08	58.89	60.82	63.35	58.87
9	−1.61	98.39	−89.22	−90.67	−85.05	−89.53
$\mathrm{VaR}_{99\%}(w; \text{one day})$			154.79	171.20	151.16	155.64
$\mathrm{ES}_{97.5\%}(w; \text{one day})$			150.04	165.10	146.37	150.84

Application to the VaR and ES

Table: Calculation of the P&L using the vega coefficient

s	S_{t+h}	Σ_{t+h}	Π_s	Π_s^{υ}	$\Pi_s^{\Delta+\upsilon}$	$\Pi_s^{\Delta+\Gamma+\upsilon}$	$\Pi_s^{\Delta+\Gamma+\Theta+\upsilon}$
1	98.07	15.58	−182.25	−79.09	−187.78	−179.71	−184.19
2	99.31	18.68	−65.61	−23.62	−62.48	−61.45	−65.92
3	99.29	16.96	−97.23	−54.40	−94.38	−93.29	−97.77
4	99.27	22.88	6.87	51.54	10.43	11.58	7.10
5	101.22	19.87	65.20	−2.33	66.38	69.61	65.13
6	101.01	19.92	53.24	−1.43	55.45	57.66	53.18
7	101.04	21.29	79.03	23.08	81.65	84.00	79.52
8	101.08	22.93	110.21	52.43	113.25	115.78	111.30
9	98.39	20.85	−74.21	15.21	−75.46	−69.84	−74.32
$\mathrm{VaR}_{99\%}(w;\text{one day})$			181.70	77.57	190.77	179.29	183.76
$\mathrm{ES}_{97.5\%}(w;\text{one day})$			172.09	73.90	184.90	169.34	173.81

The hybrid method

The hybrid method consists of combining the two approaches:

1. we first calculate the P&L for each (historical or simulated) scenario with the method based on the sensitivities;
2. we then identify the worst scenarios;
3. we finally revalue these worst scenarios by using the full pricing method.

$\Rightarrow$ The underlying idea is to consider the faster approach to locate the value-at-risk, and then to use the most accurate approach to calculate the right value

The hybrid method

Table: The 10 worst scenarios identified by the hybrid method

i	Full pricing		Greeks					
			$\Delta - \Gamma - \Theta - \upsilon$		$\Delta - \Theta$		$\Delta - \Theta - \upsilon$	
	s	Π_s	s	Π_s	s	Π_s	s	Π_s
1	100	−183.86	100	−186.15	182	−187.50	134	−202.08
2	1	−182.25	1	−184.19	169	−176.80	100	−198.22
3	134	−181.15	134	−183.34	27	−174.55	1	−192.26
4	27	−163.01	27	−164.26	134	−170.05	169	−184.32
5	169	−162.82	169	−164.02	69	−157.66	27	−184.04
6	194	−159.46	194	−160.93	108	−150.90	194	−175.36
7	49	−150.25	49	−151.43	194	−149.77	49	−165.41
8	245	−145.43	245	−146.57	49	−147.52	182	−164.96
9	182	−142.21	182	−142.06	186	−145.27	245	−153.37
10	79	−135.55	79	−136.52	100	−137.38	69	−150.68

Backtesting

mark-to-model $\neq$ mark-to-market

For on-exchange products, the simulated P&L is equal to:

$$\Pi_s(w) = \underbrace{P_{t+1}(w)}_{\text{mark-to-model}} - \underbrace{P_t(w)}_{\text{mark-to-market}}$$

whereas the realized P&L is equal to:

$$\Pi(w) = \underbrace{P_{t+1}(w)}_{\text{mark-to-market}} - \underbrace{P_t(w)}_{\text{mark-to-market}}$$

Backtesting

For exotic options and OTC derivatives, the simulated P&L is the difference between two mark-to-model values:

$$\Pi_s(w) = \underbrace{P_{t+1}(w)}_{\text{mark-to-model}} - \underbrace{P_t(w)}_{\text{mark-to-model}}$$

and the realized P&L is also the difference between two mark-to-model values:

$$\Pi(w) = \underbrace{P_{t+1}(w)}_{\text{mark-to-model}} - \underbrace{P_t(w)}_{\text{mark-to-model}}$$

$\Rightarrow$ **Model risk**

Model risk

4 types of model risk:

1. Operational risk
2. Parameter risk
3. Mis-specification risk
4. Hedging risk

(HFRM, Chapter 9, Page 491)

On the importance of risk allocation

Let us consider two trading desks A and B, whose risk measure is respectively $\mathcal{R}(w_A)$ and $\mathcal{R}(w_B)$. At the global level, the risk measure is equal to $\mathcal{R}(w_{A+B})$. The question is then how to allocate $\mathcal{R}(w_{A+B})$ to the trading desks A and B:

$$\mathcal{R}(w_{A+B}) = \mathcal{RC}_A(w_{A+B}) + \mathcal{RC}_B(w_{A+B})$$

Remark

This question is an important issue for the bank because risk allocation means capital allocation:

$$\mathcal{K}(w_{A+B}) = \mathcal{K}_A(w_{A+B}) + \mathcal{K}_B(w_{A+B})$$

Capital allocation is not neutral, because **it will impact the profitability of business units that compose the bank**

Euler allocation principle

- We decompose the P&L as follows:

$$\Pi = \sum_{i=1}^{n} \Pi_i$$

  where Π_i is the P&L of the i^{th} sub-portfolio
- We note $\mathcal{R}(\Pi)$ the risk measure associated with the P&L
- We consider the risk-adjusted performance measure (RAPM) defined by:

$$\mathrm{RAPM}(\Pi) = \frac{\mathbb{E}[\Pi]}{\mathcal{R}(\Pi)}$$

- We consider the portfolio-related RAPM of the i^{th} sub-portfolio defined by:

$$\mathrm{RAPM}(\Pi_i \mid \Pi) = \frac{\mathbb{E}[\Pi_i]}{\mathcal{R}(\Pi_i \mid \Pi)}$$

Euler allocation principle

Based on the notion of RAPM, Tasche (2008) states two properties of **risk contributions** that are desirable from an economic point of view:

1. Risk contributions $\mathcal{R}(\Pi_i \mid \Pi)$ to portfolio-wide risk $\mathcal{R}(\Pi)$ satisfy the full allocation property if:

$$\sum_{i=1}^{n} \mathcal{R}(\Pi_i \mid \Pi) = \mathcal{R}(\Pi)$$

2. Risk contributions $\mathcal{R}(\Pi_i \mid \Pi)$ are RAPM compatible if there are some $\varepsilon_i > 0$ such that:

$$\mathrm{RAPM}(\Pi_i \mid \Pi) > \mathrm{RAPM}(\Pi) \Rightarrow \mathrm{RAPM}(\Pi + h\Pi_i) > \mathrm{RAPM}(\Pi)$$

for all $0 < h < \varepsilon_i$

$\Rightarrow$ This property means that assets with a better risk-adjusted performance than the portfolio continue to have a better RAPM if their allocation increases in a small proportion

Euler allocation principle

Tasche (2008) shows that if there are risk contributions that are RAPM compatible, then $\mathcal{R}(\Pi_i \mid \Pi)$ is uniquely determined as:

$$\mathcal{R}(\Pi_i \mid \Pi) = \left. \frac{\mathrm{d}}{\mathrm{d}h} \mathcal{R}(\Pi + h\Pi_i) \right|_{h=0}$$

and the risk measure is homogeneous of degree 1

If we consider the risk measure $\mathcal{R}(w)$ defined in terms of weights, the risk contribution of sub-portfolio i is uniquely defined as:

$$\mathcal{RC}_i = w_i \frac{\partial\, \mathcal{R}(w)}{\partial\, w_i}$$

and the risk measure satisfies the Euler decomposition (or the Euler allocation principle):

$$\mathcal{R}(w) = \sum_{i=1}^{n} w_i \frac{\partial\, \mathcal{R}(w)}{\partial\, w_i} = \sum_{i=1}^{n} \mathcal{RC}_i$$

Application to Gaussian risk measures

If we assume that the portfolio return $R(w)$ is a linear function of the weights w, the expression of the standard deviation-based risk measure becomes:

$$\mathcal{R}(w) = -\mu(w) + c \cdot \sigma(w) = -w^\top \mu + c \cdot \sqrt{w^\top \Sigma w}$$

where μ and Σ are the mean vector and the covariance matrix of sub-portfolios

We have:

$$\frac{\partial\, \mathcal{R}(w)}{\partial\, w} = -\mu + c \cdot \frac{1}{2}\left(w^\top \Sigma w\right)^{-1/2}(2\Sigma w) = -\mu + c \cdot \frac{\Sigma w}{\sqrt{w^\top \Sigma w}}$$

The risk contribution of the i^{th} sub-portfolio is then:

$$\mathcal{RC}_i = w_i \cdot \left(-\mu_i + c \cdot \frac{(\Sigma w)_i}{\sqrt{w^\top \Sigma w}}\right)$$

Application to Gaussian risk measures

We verify that the standard deviation-based risk measure satisfies the full allocation property:

$$\begin{aligned}
\sum_{i=1}^{n} \mathcal{RC}_i &= \sum_{i=1}^{n} w_i \cdot \left(-\mu_i + c \cdot \frac{(\Sigma w)_i}{\sqrt{w^\top \Sigma w}} \right) \\
&= w^\top \left(-\mu + c \cdot \frac{\Sigma w}{\sqrt{w^\top \Sigma w}} \right) \\
&= -w^\top \mu + c \cdot \sqrt{w^\top \Sigma w} \\
&= \mathcal{R}(w)
\end{aligned}$$

Application to Gaussian risk measures

- Gaussian VaR risk contribution:

$$\mathcal{RC}_i = w_i \cdot \left(-\mu_i + \Phi^{-1}(\alpha) \cdot \frac{(\Sigma w)_i}{\sqrt{w^\top \Sigma w}} \right)$$

- Gaussian ES risk contribution:

$$\mathcal{RC}_i = w_i \cdot \left(-\mu_i + \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{(1-\alpha)} \cdot \frac{(\Sigma w)_i}{\sqrt{w^\top \Sigma w}} \right)$$

Application to Gaussian risk measures

block

We consider the Apple/Coca-Cola portfolio that has been used for calculating the Gaussian VaR. We recall that the nominal exposures were \$1 093.3 (Apple) and \$842.8 (Coca-Cola), the estimated standard deviation of daily returns was equal to 1.3611% for Apple and 0.9468% for Coca-Cola and the cross-correlation of stock returns was equal to 12.0787%.

Application to Gaussian risk measures

Table: Risk decomposition of the 99% Gaussian value-at-risk

Asset	w_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
Apple	1093.3	2.83%	30.96	75.14%
Coca-Cola	842.8	1.22%	10.25	24.86%
$\mathcal{R}(w)$			41.21	

Table: Risk decomposition of the 99% Gaussian expected shortfall

Asset	w_i	$\mathcal{MR}_i$	$\mathcal{RC}_i$	$\mathcal{RC}_i^\star$
Apple	1093.3	3.24%	35.47	75.14%
Coca-Cola	842.8	1.39%	11.74	24.86%
$\mathcal{R}(w)$			47.21	

Application to non-normal risk measures

Generalized formulas

- The risk contribution for the value-at-risk is equal to:

$$\mathcal{RC}_i = \mathbb{E}\left[L_i \mid L(w) = \mathrm{VaR}_\alpha(L)\right]$$

- The risk contribution for the expected shortfall is equal to:

$$\mathcal{RC}_i = \mathbb{E}\left[L_i \mid L(w) \geq \mathrm{VaR}_\alpha(L)\right]$$

$\Rightarrow$ These formulas can easily be applied to historical and Monte Carlo risk measures (HFRM, pages 109-116)

Calculating the Gaussian VaR risk contribution

Asset returns are assumed to be Gaussian:

$$R \sim \mathcal{N}(\mu, \Sigma)$$

The portfolio's loss is equal to:

$$L(w) = -R(w) = -\sum_{i=1}^{n} w_i R_i = -w^\top R$$

We notice that:

$$L_i = -w_i R_i$$

and:

$$\mathbb{E}[L_i \mid L(w) = \mathrm{VaR}_\alpha(w;h)] = -w_i \mathbb{E}[R_i \mid L(w) = \mathrm{VaR}_\alpha(w;h)]$$

Calculating the Gaussian VaR risk contribution

We have:

$$\left(\begin{array}{c} R \\ L(w) \end{array}\right) = \left(\begin{array}{c} I_n \\ -w^\top \end{array}\right) R$$

and:

$$\left(\begin{array}{c} R \\ L(w) \end{array}\right) \sim \mathcal{N}\left(\left(\begin{array}{c} \mu \\ -w^\top \mu \end{array}\right), \left(\begin{array}{cc} \Sigma & -\Sigma w \\ -w^\top \Sigma & w^\top \Sigma w \end{array}\right)\right)$$

We would like to calculate:

$$\mathcal{RC}_i = -w_i \mathbb{E}\left[R_i \mid L(w) = \mathrm{VaR}_\alpha(w; h)\right]$$

Conditional distribution in the case of the normal distribution

Let us consider a Gaussian random vector defined as follows:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{x,x} & \Sigma_{x,y} \\ \Sigma_{y,x} & \Sigma_{y,y} \end{pmatrix}\right)$$

We have:

$$Y \mid X = x \sim \mathcal{N}\left(\mu_{y|x}, \Sigma_{y,y|x}\right)$$

where:

$$\mu_{y|x} = \mathbb{E}\left[Y \mid X = x\right] = \mu_y + \Sigma_{y,x}\Sigma_{x,x}^{-1}\left(x - \mu_x\right)$$

and:

$$\Sigma_{y,y|x} = \operatorname{cov}\left(Y \mid X = x\right) = \Sigma_{y,y} - \Sigma_{y,x}\Sigma_{x,x}^{-1}\Sigma_{x,y}$$

Calculating the Gaussian VaR risk contribution

Since $\mathrm{VaR}_\alpha(w;h) = -w^\top \mu + \Phi^{-1}(\alpha)\sqrt{w^\top \Sigma w}$, we have:

$$\begin{aligned}
\mathbb{E}\left[R \mid L(w) = \mathrm{VaR}_\alpha(w;h)\right] &= \mathbb{E}\left[R \mid L(w) = -w^\top \mu + \Phi^{-1}(\alpha)\sqrt{w^\top \Sigma w}\right] \\
&= \mu - \Sigma w \left(w^\top \Sigma w\right)^{-1} \cdot \\
&\quad \left(-w^\top \mu + \Phi^{-1}(\alpha)\sqrt{w^\top \Sigma w} - \left(-w^\top \mu\right)\right) \\
&= \mu - \Phi^{-1}(\alpha)\Sigma w \frac{\sqrt{w^\top \Sigma w}}{\left(w^\top \Sigma w\right)^{-1}} \\
&= \mu - \Phi^{-1}(\alpha)\frac{\Sigma w}{\sqrt{w^\top \Sigma w}}
\end{aligned}$$

and:

$$\mathcal{RC}_i = -w_i\left(\mu - \Phi^{-1}(\alpha)\frac{\Sigma w}{\sqrt{w^\top \Sigma w}}\right)_i = -w_i \mu_i + \Phi^{-1}(\alpha)\frac{w_i \cdot (\Sigma w)_i}{\sqrt{w^\top \Sigma w}}$$

Exercises

- Value-at-risk
  - Exercise 2.4.2 – Covariance matrix
  - Exercise 2.4.4 – Value-at-risk of a long/short portfolio
  - Exercise 2.4.4 – Value-at-risk of an equity portfolio hedged with put options
- Expected shortfall
  - Exercise 2.4.10 – Expected shortfall of an equity portfolio
  - Exercise 2.4.11 – Risk measure of a long/short portfolio
- Options and derivatives
  - Exercise 2.4.6 – Risk management of exotic options
  - Exercise 2.4.7 – P&L approximation with Greek sensitivities

References

- Basel Committee on Banking Supervision (1996)
  Amendment to the Capital Accord to Incorporate Market Risks, January 1996
- Basel Committee on Banking Supervision (2009
  Revisions to the Basel II Market Risk Framework, July 2009
- Basel Committee on Banking Supervision (2019)
  Minimum Capital Requirements for Market Risk, January 2019.
- RONCALLI, T. (2020)
  Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 2.

Course 2023-2024 in Financial Risk Management
Lecture 3. Credit Risk

Thierry Roncalli*

*Amundi Asset Management[11]

*University of Paris-Saclay

September 2023

[11] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

The loan market

$\Rightarrow$ Banking intermediation (retail banks and corporate investment banks) $\neq$ financial market of debt securities (money market, bonds, notes, etc.)

Counterparties

- Sovereign
- Financial
- Corporate
- Retail

Products

- Mortgage and housing debt, consumer credit (auto loans, credit cards, revolving credit), student loans
- Revolving credit facilities (for corporates), corporate loans and other credit lines

$\Rightarrow$ Differences in terms of products and maturities (retail $\neq$ corporate)

Credit decision process

- Segmentation (retail banking)
- Pricing of the credit spread (commercial and investment banking)

The loan market

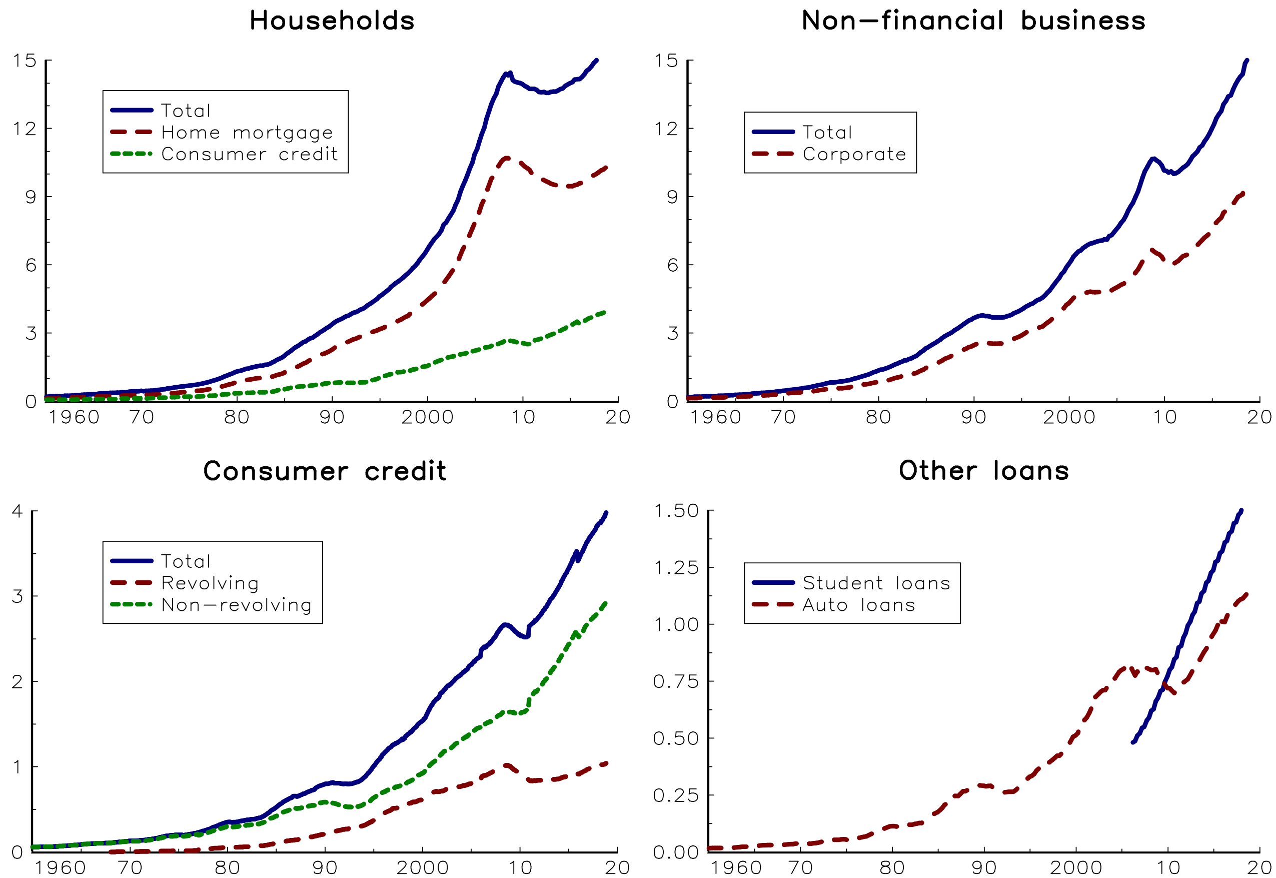


Figure: Credit debt outstanding in the United States (in $ tn)

The loan market

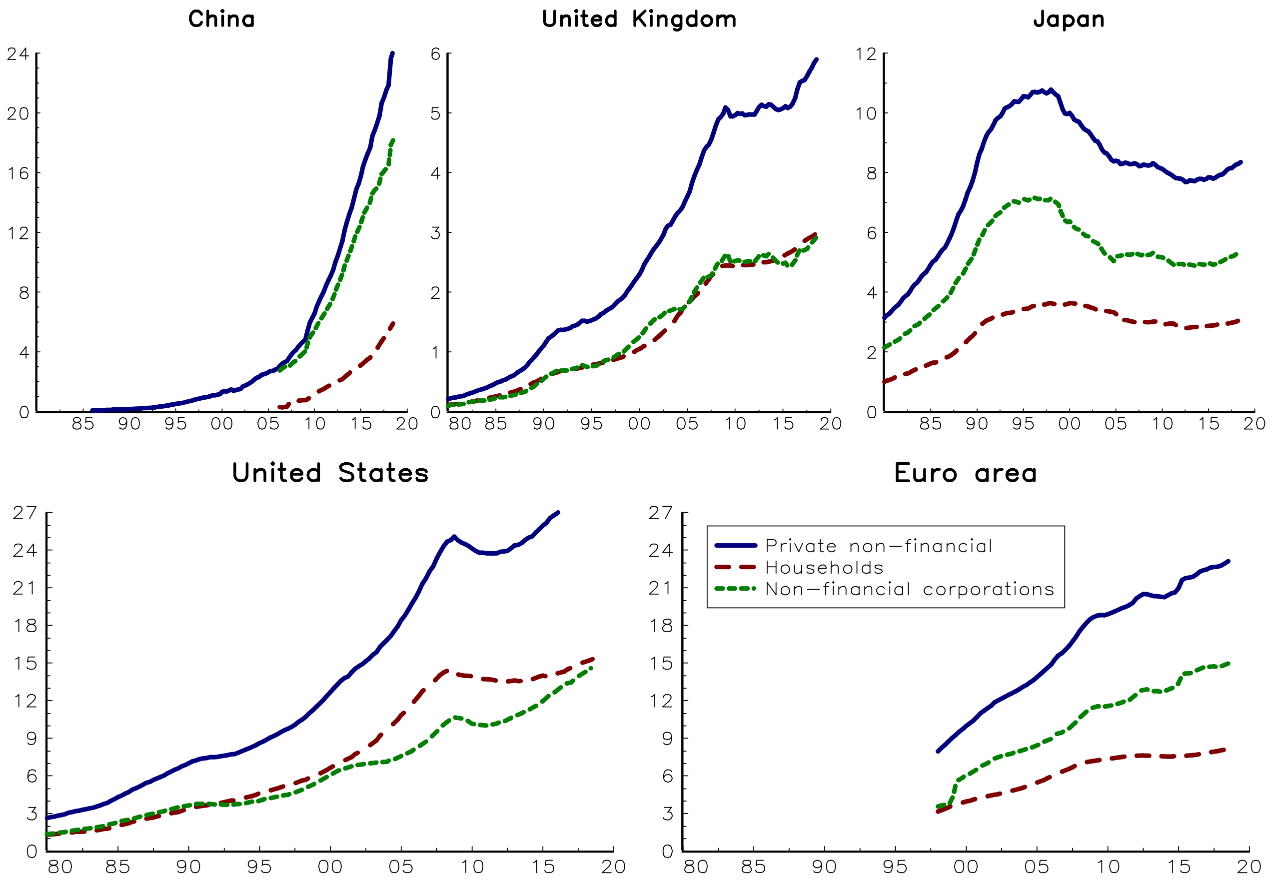


Figure: Credit to the private non-financial sector (in $ tn)

The bond market

Issuance $\neq$ outstanding:

- Primary market
- Secondary market

Three main sectors

- Central and local governments
- Financials
- Corporates

Statistics of the bond market

Table: Debt securities by residence of issuer (in $ bn)

		Dec. 2004	Dec. 2007	Dec. 2010	Dec. 2017
Canada	Gov.	682	841	1 149	1 264
	Fin.	283	450	384	655
	Corp.	212	248	326	477
	Total	1 180	1 544	1 863	2 400
France	Gov.	1 236	1 514	1 838	2 258
	Fin.	968	1 619	1 817	1 618
	Corp.	373	382	483	722
	Total	2 576	3 515	4 138	4 597
Germany	Gov.	1 380	1 717	2 040	1 939
	Fin.	2 296	2 766	2 283	1 550
	Corp.	133	174	168	222
	Total	3 809	4 657	4 491	3 712
Italy	Gov.	1 637	1 928	2 069	2 292
	Fin.	772	1 156	1 403	834
	Corp.	68	95	121	174
	Total	2 477	3 178	3 593	3 299

Statistics of the bond market

Table: Debt securities by residence of issuer (in $ bn)

		Dec. 2004	Dec. 2007	Dec. 2010	Dec. 2017
Japan	Gov.	6 336	6 315	10 173	9 477
	Fin.	2 548	2 775	3 451	2 475
	Corp.	1 012	762	980	742
	Total	9 896	9 852	14 604	12 694
Spain	Gov.	462	498	796	1 186
	Fin.	434	1 385	1 442	785
	Corp.	15	19	19	44
	Total	910	1 901	2 256	2 015
UK	Gov.	798	1 070	1 674	2 785
	Fin.	1 775	3 127	3 061	2 689
	Corp.	452	506	473	533
	Total	3 027	4 706	5 210	6 011
US	Gov.	6 459	7 487	12 072	17 592
	Fin.	12 706	17 604	15 666	15 557
	Corp.	3 004	3 348	3 951	6 137
	Total	22 371	28 695	31 960	39 504

Statistics of the bond market

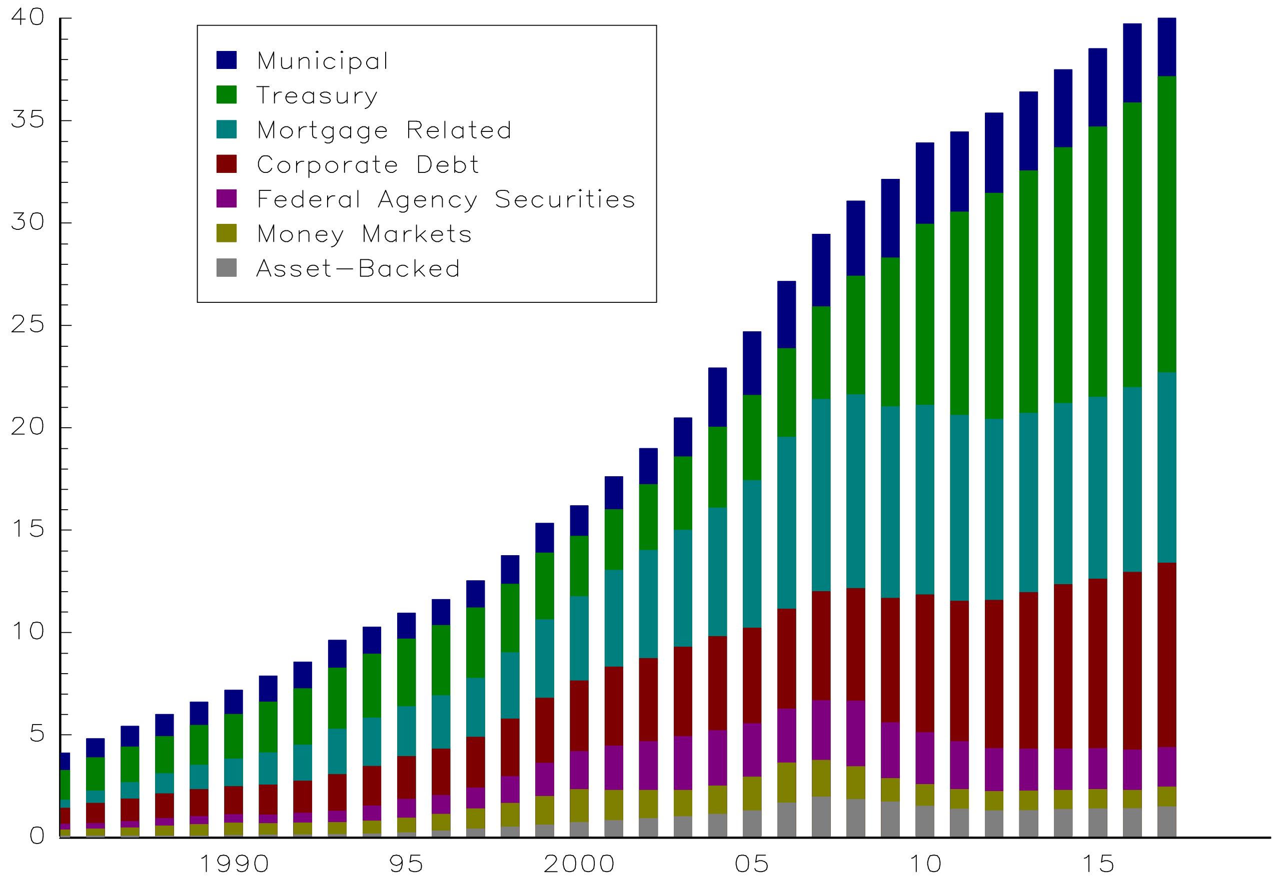


Figure: US bond market outstanding (in $ tn)

Statistics of the bond market

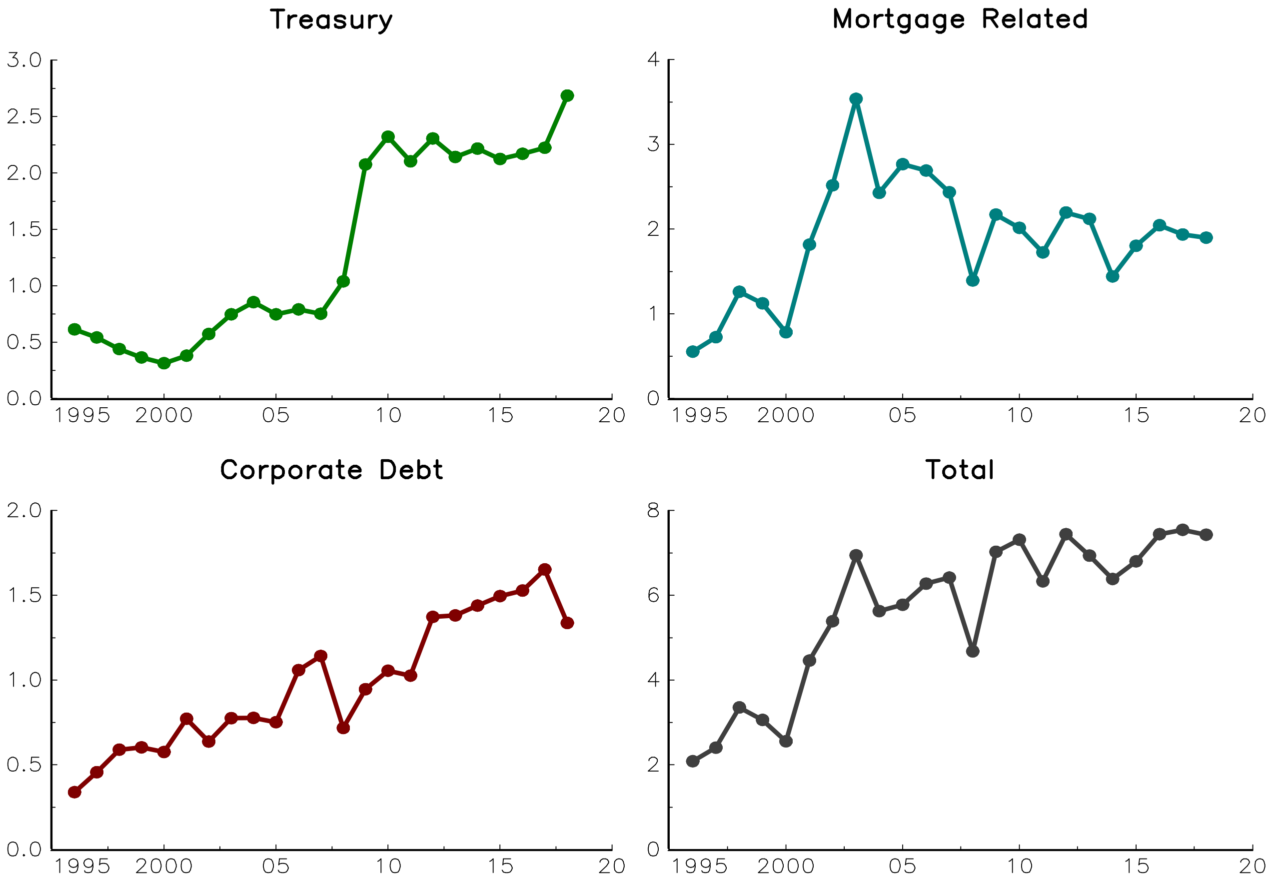


Figure: US bond market issuance (in $ tn)

Statistics of the bond market

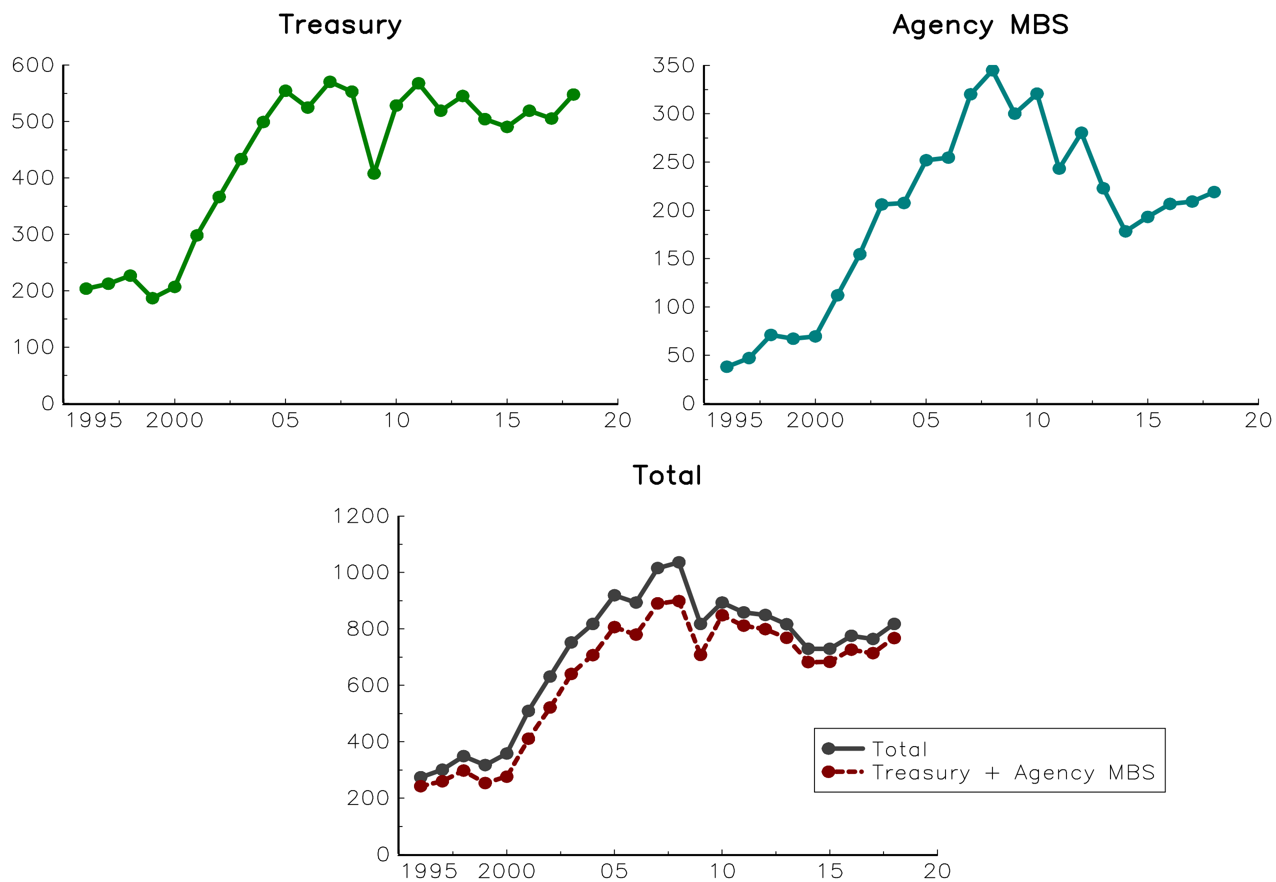


Figure: Average daily trading volume in US bond markets (in $ bn)

Bond pricing (without default risk)

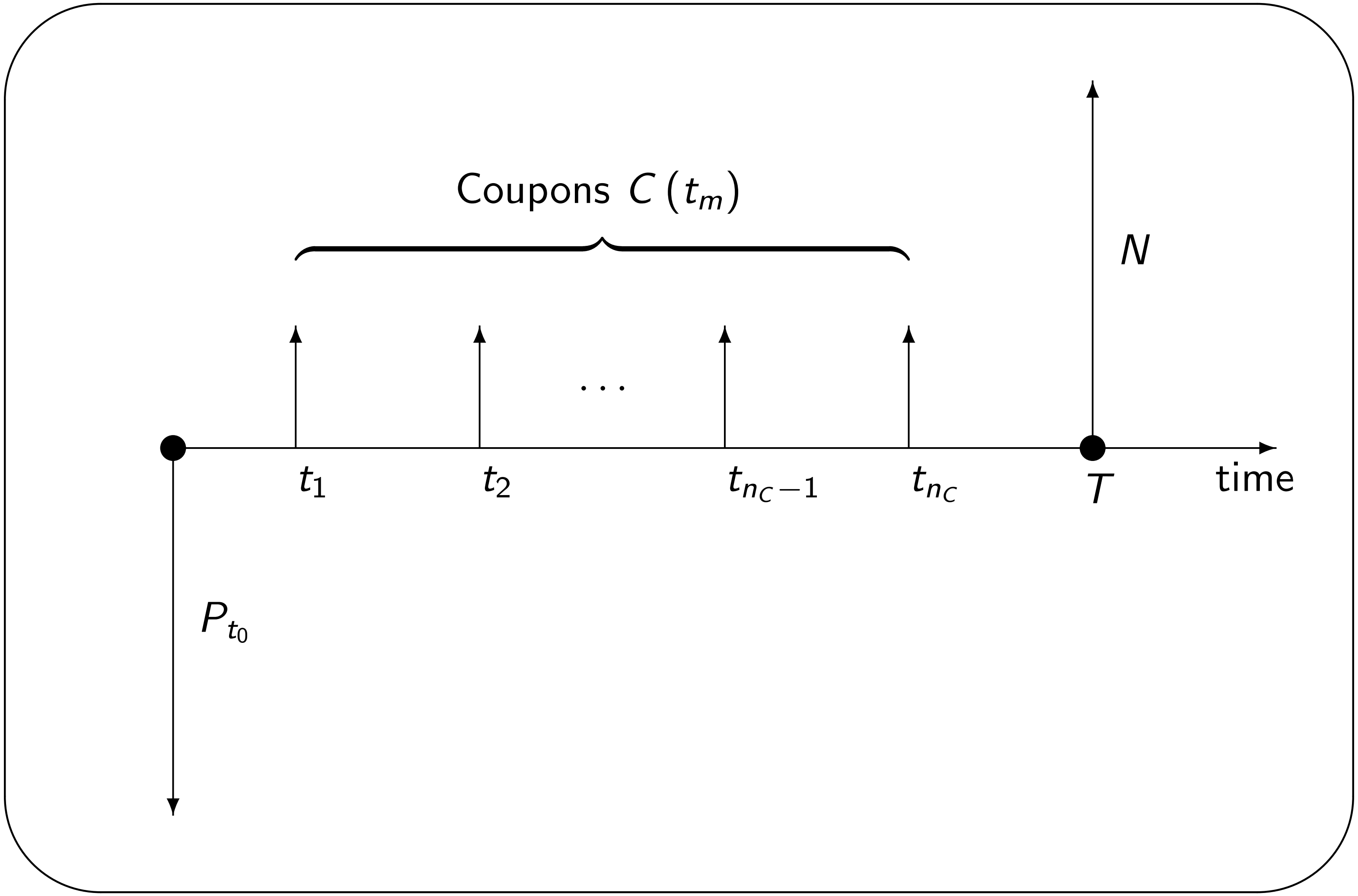


Figure: Cash flows of a bond with a fixed coupon rate

Bond pricing (without default risk)

The price of the bond at the inception date t_0 is the sum of the present values of all the expected coupon payments and the par value:

$$P_{t_0} = \sum_{m=1}^{n_C} C(t_m) \cdot B_{t_0}(t_m) + N \cdot B_{t_0}(T)$$

where $B_t(t_m)$ is the discount factor at time t for the maturity date t_m

Bond pricing (without default risk)

If we take into account the accrued interests, we have:

$$P_t + AC_t = \sum_{t_m \geq t} C(t_m) \cdot B_t(t_m) + N \cdot B_t(T)$$

Here, AC_t is the accrued coupon:

$$AC_t = C(t_c) \cdot \frac{t - t_c}{365}$$

and t_c is the last coupon payment date with $c = \{m : t_{m+1} > t, t_m \leq t\}$

- $P_t + AC_t$ is called the '*dirty price*'
- P_t is called the '*clean price*'

Impact of the term structure

3 main movements:

1. The movement of level corresponds to a parallel shift of interest rates.
2. A twist in the slope of the yield curve indicates how the spread between long and short interest rates moves.
3. A change in the curvature of the yield curve affects the convexity of the term structure.

Impact of the term structure

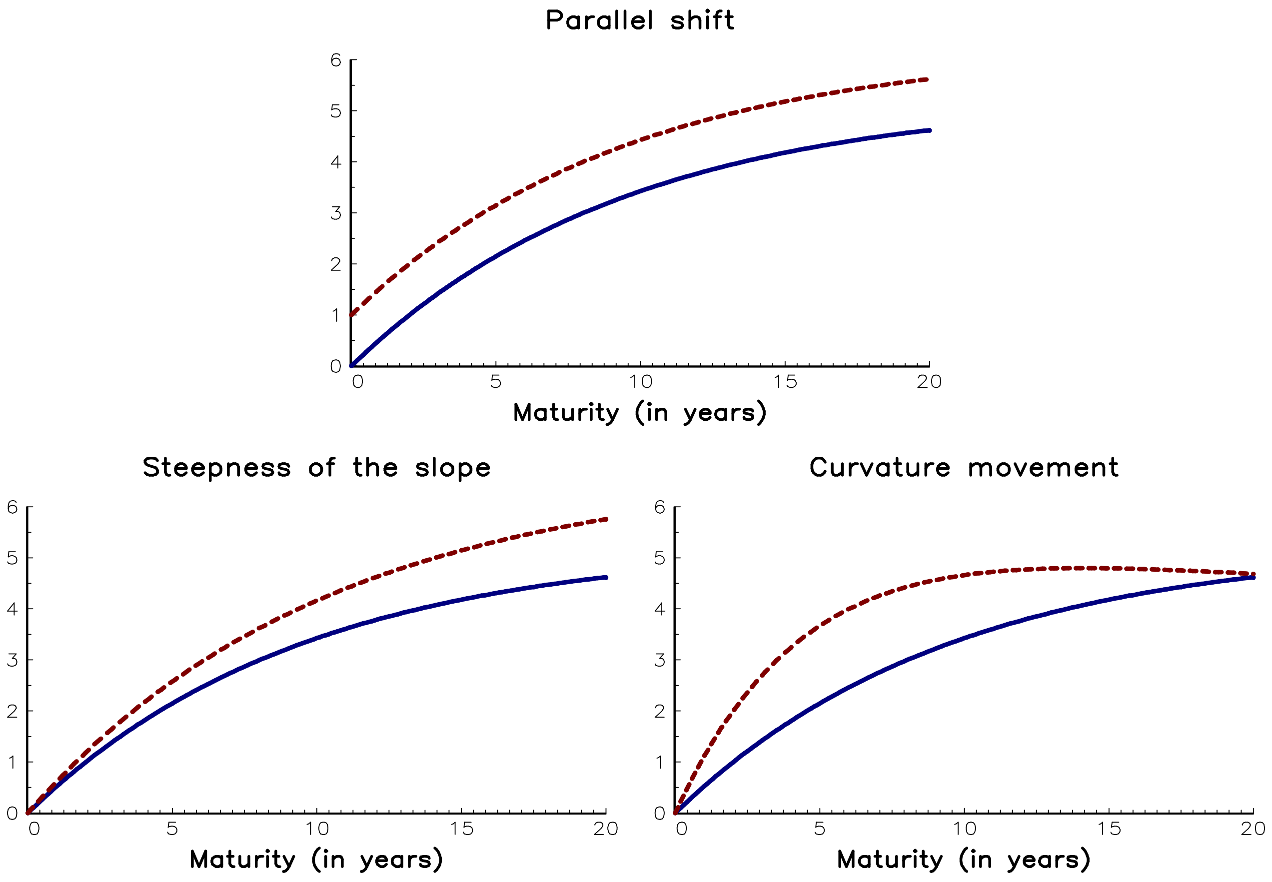


Figure: Movements of the yield curve

Yield to maturity

The yield to maturity y of a bond is the constant discount rate which returns its market price:

$$\sum_{t_m \geq t} C(t_m) e^{-(t_m - t)y} + N e^{-(T-t)y} = P_t + AC_t$$

The sensitivity S is the derivative of the clean price P_t with respect to the yield to maturity y:

$$S = \frac{\partial P_t}{\partial y} = -\sum_{t_m \geq t} (t_m - t) C(t_m) e^{-(t_m - t)y} - (T - t) N e^{-(T-t)y}$$

$\Rightarrow$ It indicates how the P&L of a long position on the bond moves when the yield to maturity changes:

$$\Pi \approx S \cdot \Delta y$$

Yield to maturity

Example

We assume that the zero-coupon rates are equal to 0.52% (1Y), 0.99% (2Y), 1.42% (3Y), 1.80% (4Y) and 2.15% (5Y). We consider a bond with a constant annual coupon of 5%. The nominal of the bond is \$100. We would like to price the bond when the maturity T ranges from 1 to 5 years.

The price of the four-year bond is equal to:

$$P_t = \frac{5}{(1+0.52\%)} + \frac{5}{(1+0.99\%)^2} + \frac{5}{(1+1.42\%)^3} + \frac{105}{(1+1.80\%)^4} = \$112.36$$

Yield to maturity

Table: Price, yield to maturity and sensitivity of bonds

T	$R_t(T)$	$B_t(T)$	P_t	y	S
1	0.52%	99.48	104.45	0.52%	−104.45
2	0.99%	98.03	107.91	0.98%	−210.86
3	1.42%	95.83	110.50	1.39%	−316.77
4	1.80%	93.04	112.36	1.76%	−420.32
5	2.15%	89.82	113.63	2.08%	−520.16

Yield to maturity

Table: Impact of a parallel shift of the yield curve on the bond with five-year maturity

ΔR (in bps)	$\check{P}_t$	ΔP_t	$\hat{P}_t$	ΔP_t	$S \times \Delta y$
-50	116.26	2.63	116.26	2.63	2.60
-30	115.20	1.57	115.20	1.57	1.56
-10	114.15	0.52	114.15	0.52	0.52
0	113.63	0.00	113.63	0.00	0.00
10	113.11	-0.52	113.11	-0.52	-0.52
30	112.08	-1.55	112.08	-1.55	-1.56
50	111.06	-2.57	111.06	-2.57	-2.60

$$\check{P}_t = \sum_{t_m \geq t} C(t_m) e^{-(t_m - t)(R_t(t_m) + \Delta R)} + N e^{-(T-t)(R_t(T) + \Delta R)}$$

$$\hat{P}_t = \sum_{t_m \geq t} C(t_m) e^{-(t_m - t)\left(y + \Delta R\right)} + N e^{-(T-t)\left(y + \Delta R\right)}$$

Bond pricing (with default risk)

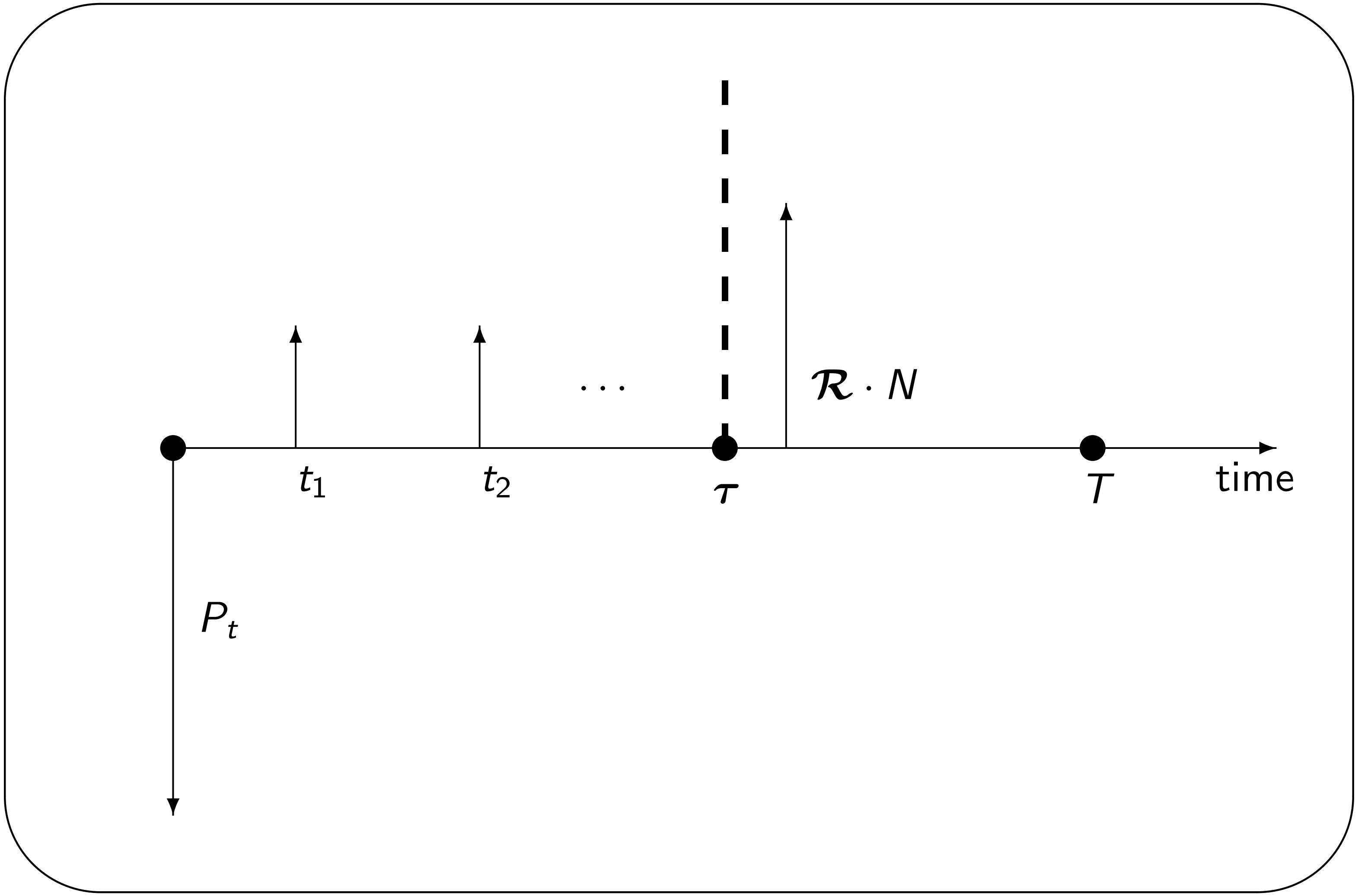


Figure: Cash flows of a bond with default risk

Bond pricing (with default risk)

- the coupons $C(t_m)$ if the bond issuer does not default before the coupon date t_m:

$$\sum_{t_m \geq t} C(t_m) \cdot \mathbb{1}\{\boldsymbol{\tau} > t_m\}$$

- the notional if the bond issuer does not default before the maturity date:

$$N \cdot \mathbb{1}\{\boldsymbol{\tau} > T\}$$

- the recovery part if the bond issuer defaults before the maturity date:

$$\boldsymbol{\mathcal{R}} \cdot N \cdot \mathbb{1}\{\boldsymbol{\tau} \leq T\}$$

where $\boldsymbol{\mathcal{R}}$ is the corresponding recovery rate

$$\begin{aligned} SV_t &= \sum_{t_m \geq t} C(t_m) \cdot e^{-\int_t^{t_m} r_s \,\mathrm{d}s} \cdot \mathbb{1}\{\boldsymbol{\tau} > t_m\} + \\ & \quad N \cdot e^{-\int_t^{T} r_s \,\mathrm{d}s} \cdot \mathbb{1}\{\boldsymbol{\tau} > T\} + \boldsymbol{\mathcal{R}} \cdot N \cdot e^{-\int_t^{\boldsymbol{\tau}} r_s \,\mathrm{d}s} \cdot \mathbb{1}\{\boldsymbol{\tau} \leq T\} \end{aligned}$$

Bond pricing (with default risk)

Closed-form formula

$$\begin{aligned} P_t + AC_t &= \sum_{t_m \geq t} C(t_m) B_t(t_m) \mathbf{S}_t(t_m) + N B_t(T) \mathbf{S}_t(T) + \\ & \mathcal{R} N \int_t^T B_t(u) f_t(u) \,\mathrm{d}u \end{aligned}$$

where $\mathbf{S}_t(u)$ is the survival function at time u and $f_t(u)$ the associated density function

Bond pricing (with default risk)

If we consider an exponential default time with parameter λ – $\boldsymbol{\tau} \sim \mathcal{E}(\lambda)$, we have $\mathbf{S}_t(u) = e^{-\lambda(u-t)}$, $f_t(u) = \lambda e^{-\lambda(u-t)}$ and:

$$\begin{aligned} P_t + AC_t &= \sum_{t_m \geq t} C(t_m) B_t(t_m) e^{-\lambda(t_m - t)} + N B_t(T) e^{-\lambda(T-t)} + \\ & \lambda \mathcal{R} N \int_t^T B_t(u) e^{-\lambda(u-t)} \,\mathrm{d}u \end{aligned}$$

If we assume a flat yield curve – $R_t(u) = r$, we obtain:

$$\begin{aligned} P_t + AC_t &= \sum_{t_m \geq t} C(t_m) e^{-(r+\lambda)(t_m - t)} + N e^{-(r+\lambda)(T-t)} + \\ & \lambda \mathcal{R} N \left(\frac{1 - e^{-(r+\lambda)(T-t)}}{r + \lambda} \right) \end{aligned}$$

If the recovery rate is equal to zero, $y = r + \lambda$

Credit spread

The credit spread is equal to the difference between the yield to maturity with default risk y and the yield to maturity without default risk $y^{\star}$:

$$s = y - y^{\star}$$

Remark

In the previous case (exponential default time + flat yield curve + zero recovery), we have:

$$s = \lambda$$

If λ is relatively small (less than 20%), the credit spread is approximately equal to the annual default probability PD:

$$\mathrm{PD} = \mathbf{S}_t(t+1) = 1 - e^{-\lambda} \approx \lambda$$

Credit spread

We consider the previous example with a coupon of 4.5% and a 10-year maturity

Table: Computation of the credit spread s

$\mathcal{R}$ (in %)	λ (in bps)	PD (in bps)	P_t (in \$)	y (in %)	s (in bps)
0	0	0.0	110.1	3.24	0.0
	10	10.0	109.2	3.34	9.9
	200	198.0	93.5	5.22	198.1
	1000	951.6	50.4	13.13	988.9
40	0	0.0	110.1	3.24	0.0
	10	10.0	109.6	3.30	6.0
	200	198.0	99.9	4.41	117.1
	1000	951.6	73.3	8.23	498.8
80	0	0.0	110.1	3.24	0.0
	10	10.0	109.9	3.26	2.2
	200	198.0	106.4	3.66	41.7
	1000	951.6	96.3	4.85	161.4

Credit risk versus market risk

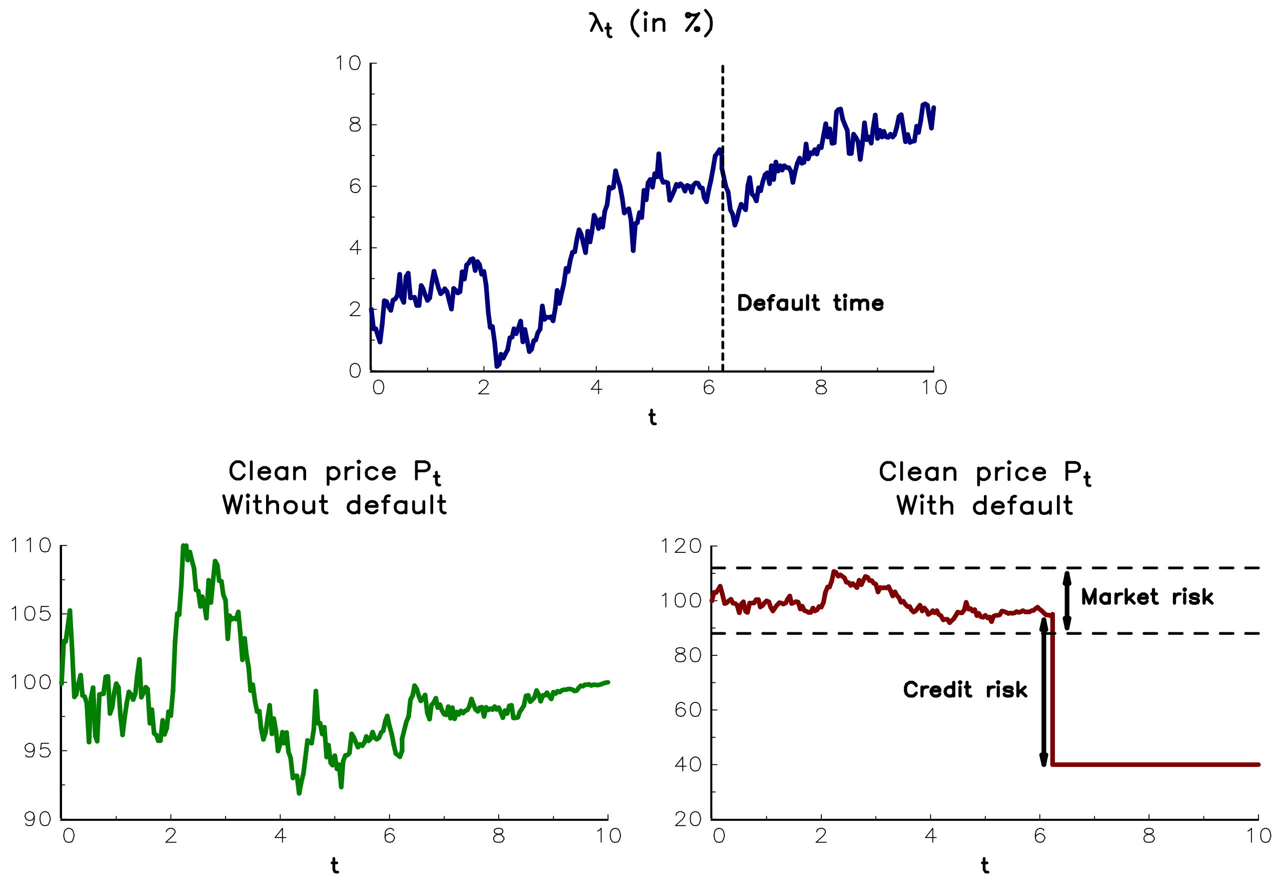


Figure: Difference between market and credit risks for a bond

Credit securitization

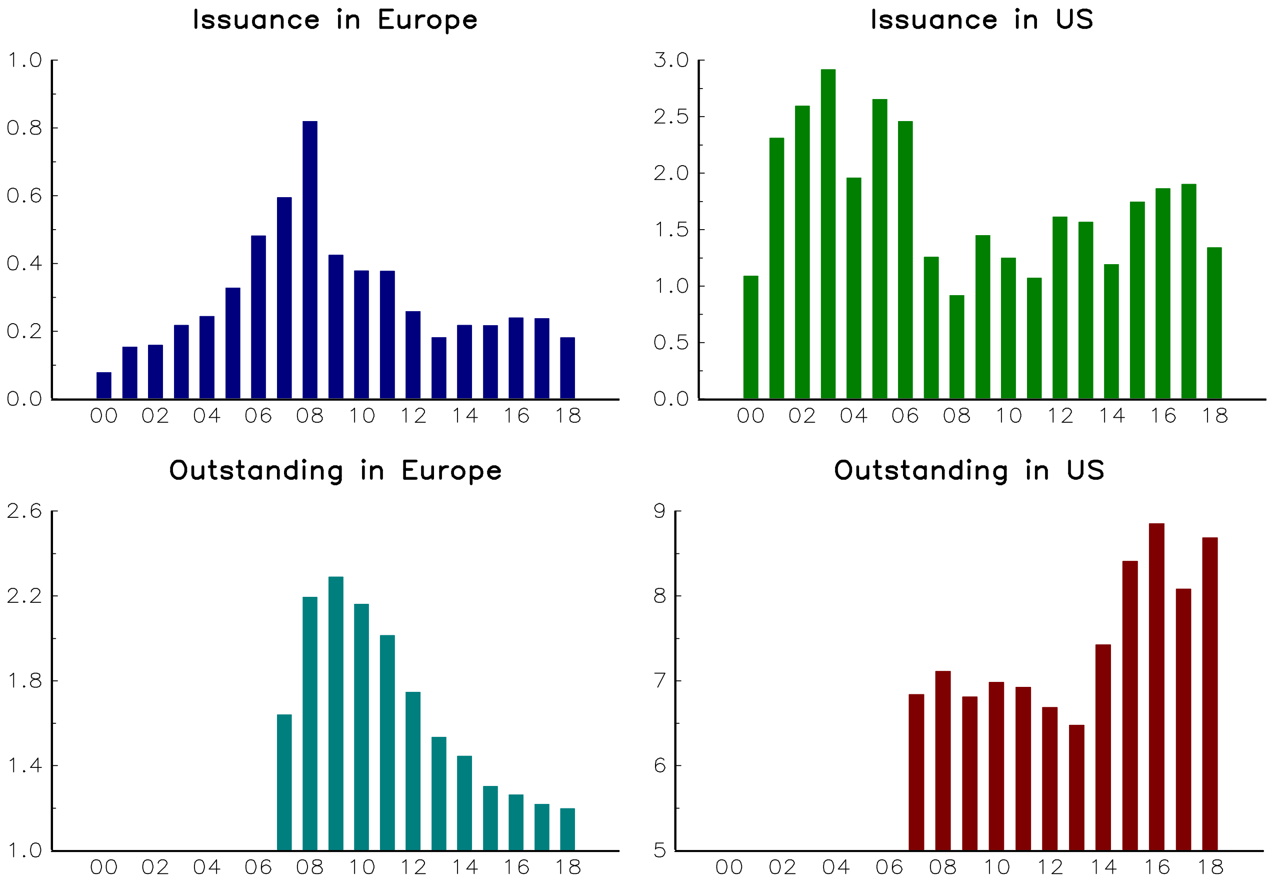


Figure: Securitization in Europe and US (in € tn)

Credit securitization

Collateral assets

- Mortgage-backed securities (MBS)
  - Residential mortgage-backed securities (RMBS)
  - Commercial mortgage-backed securities (CMBS)
- Collateralized debt obligations (CDO)
  - Collateralized loan obligations (CLO)
  - Collateralized bond obligations (CBO)
- Asset-backed securities (ABS)
  - Auto loans
  - Credit cards and revolving credit
  - Student loans

Credit securitization

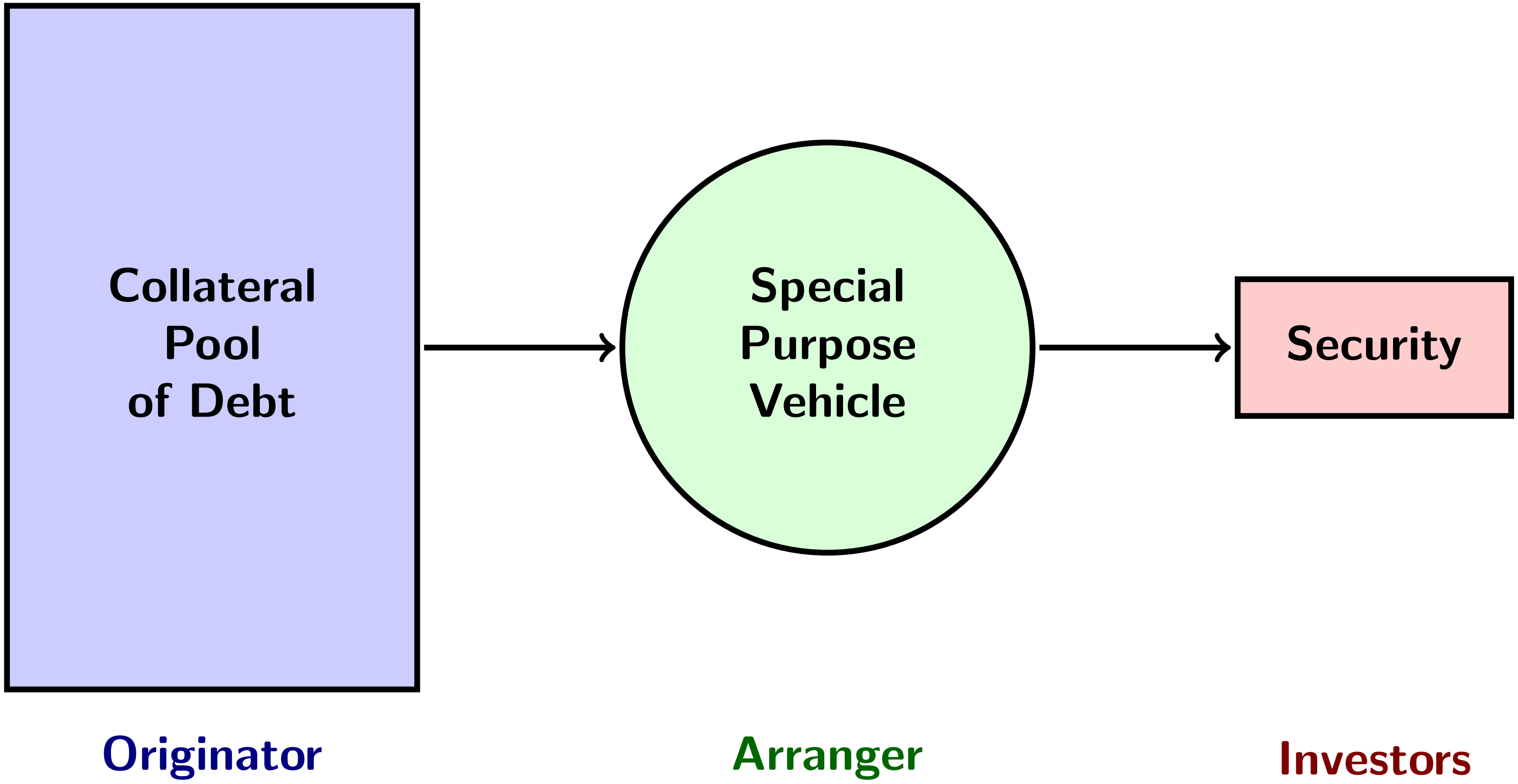


Figure: Structure of pass-through securities

Credit securitization

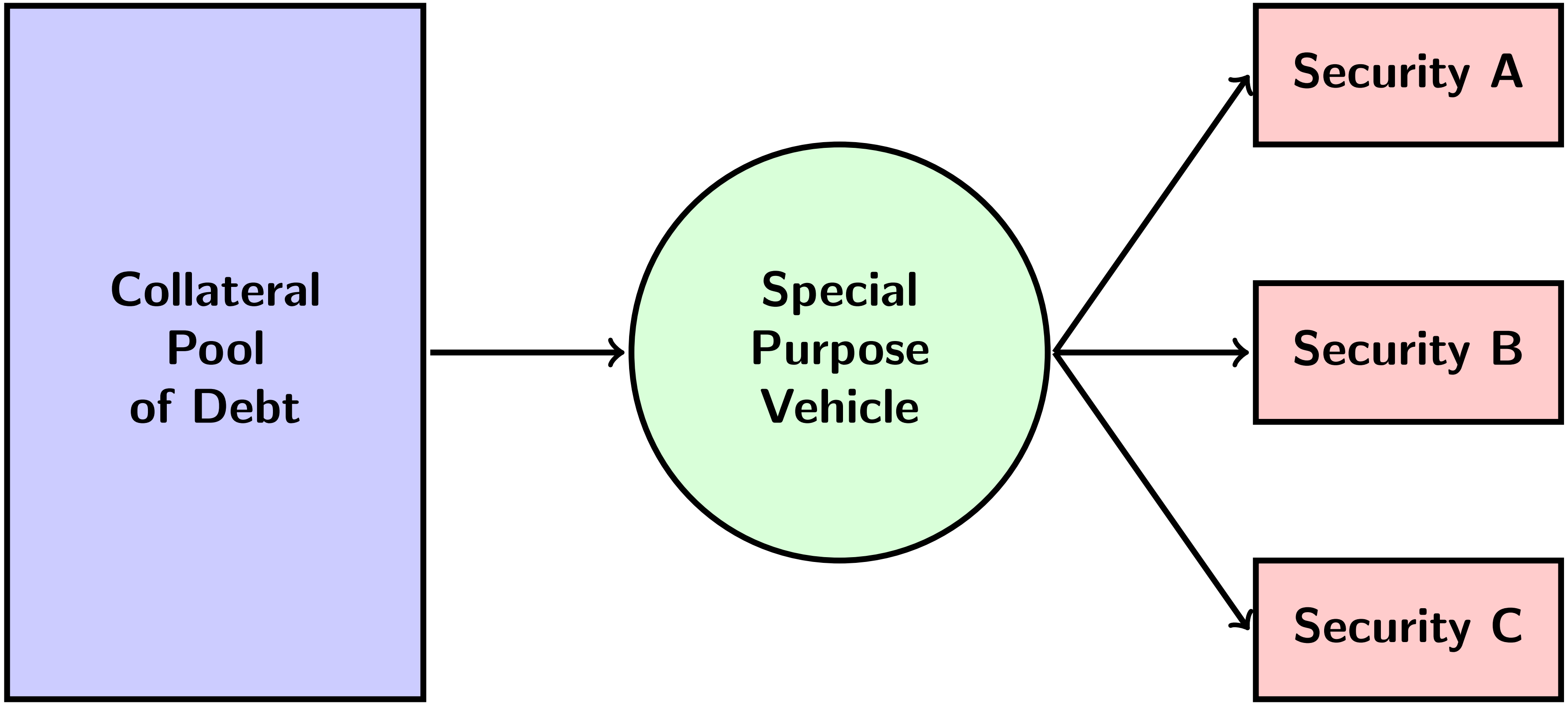


Figure: Structure of pay-through securities

Credit securitization

Table: US mortgage-backed securities

Year	Agency MBS	Agency CMO	Non-agency CMBS	Non-agency RMBS	Total (in $ bn)
Issuance					
2002	57.5%	23.6%	2.2%	16.7%	2 515
2006	33.6%	11.0%	7.9%	47.5%	2 691
2008	84.2%	10.8%	1.2%	3.8%	1 394
2010	71.0%	24.5%	1.2%	3.3%	2 013
2012	80.1%	16.4%	2.2%	1.3%	2 195
2014	68.7%	19.2%	7.0%	5.1%	1 440
2016	76.3%	15.7%	3.8%	4.2%	2 044
2018	69.2%	16.6%	4.7%	9.5%	1 899
Outstanding amount					
2002	59.7%	17.4%	5.6%	17.2%	5 289
2006	45.7%	14.9%	8.3%	31.0%	8 390
2008	52.4%	14.0%	8.8%	24.9%	9 467
2010	59.2%	14.6%	8.1%	18.1%	9 258
2012	64.0%	14.8%	7.2%	14.0%	8 838
2014	68.0%	13.7%	7.1%	11.2%	8 842
2016	72.4%	12.3%	5.9%	9.5%	9 023
2018	74.7%	11.3%	5.6%	8.4%	9 732

Credit securitization

Table: US asset-backed securities

Year	Auto Loans	CDO & CLO	Credit Cards	Equip-ement	Other	Student Loans	Total (in $ bn)
				Issuance			
2002	34.9%	21.0%	25.2%	2.6%	6.8%	9.5%	269
2006	13.5%	60.1%	9.3%	2.2%	4.6%	10.3%	658
2008	16.5%	37.8%	25.9%	1.3%	5.4%	13.1%	215
2010	46.9%	6.4%	5.2%	7.0%	22.3%	12.3%	126
2012	33.9%	23.1%	12.5%	7.1%	13.7%	9.8%	259
2014	25.2%	35.6%	13.1%	5.2%	17.0%	4.0%	393
2016	28.3%	36.8%	8.3%	4.6%	16.9%	5.1%	325
2018	20.8%	54.3%	6.1%	5.1%	10.1%	3.7%	517
				Outstanding amount			
2002	20.7%	28.6%	32.5%	4.1%	7.5%	6.6%	905
2006	11.8%	49.3%	17.6%	3.1%	6.0%	12.1%	1 657
2008	7.7%	53.5%	17.3%	2.4%	6.2%	13.0%	1 830
2010	7.6%	52.4%	14.4%	2.4%	7.1%	16.1%	1 508
2012	11.0%	48.7%	10.0%	3.3%	8.7%	18.4%	1 280
2014	13.2%	46.8%	10.1%	3.9%	9.8%	16.2%	1 349
2016	13.9%	48.0%	9.3%	3.7%	11.6%	13.5%	1 397
2018	13.3%	48.2%	7.4%	5.0%	16.0%	10.2%	1 677

Credit default swap

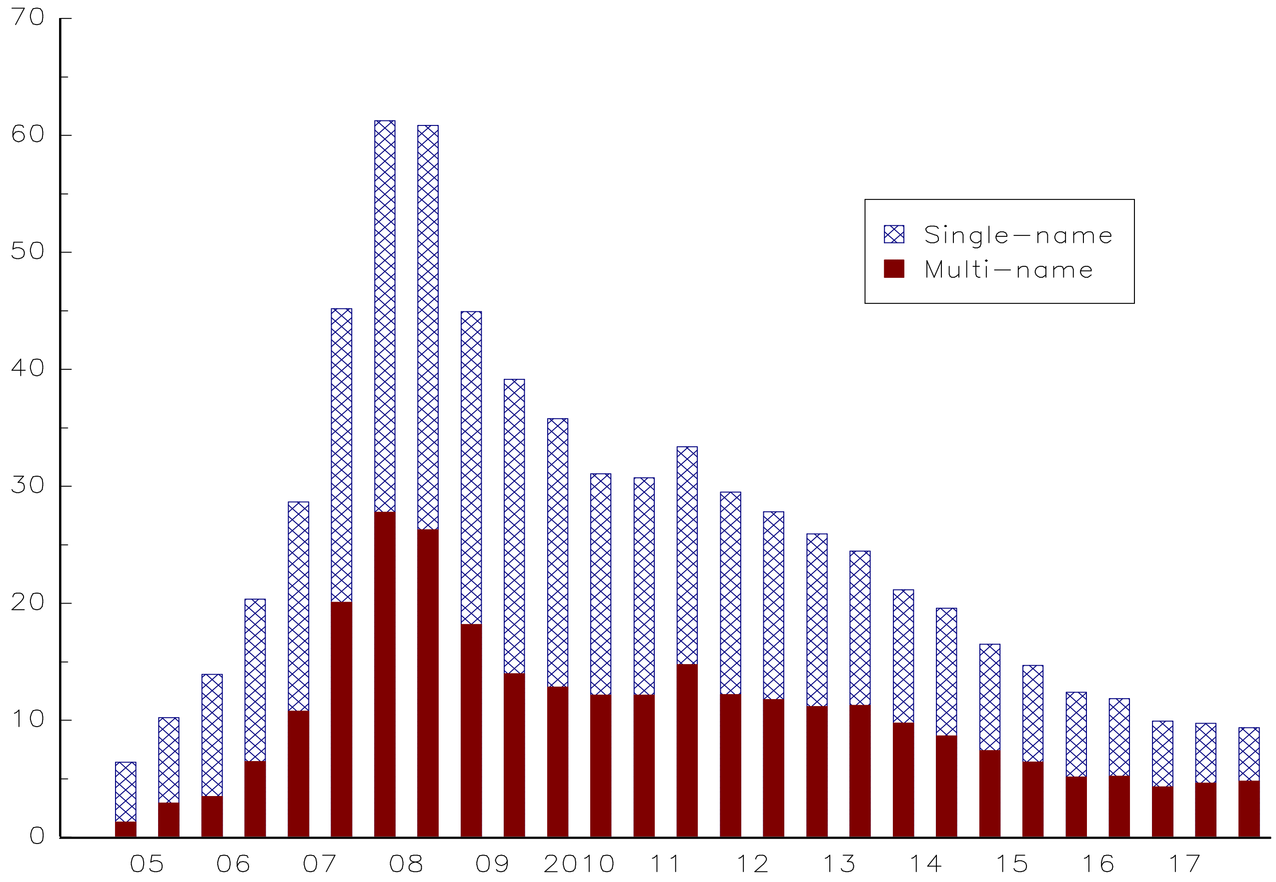


Figure: Outstanding amount of credit default swaps (in $ tn)

Credit default swap

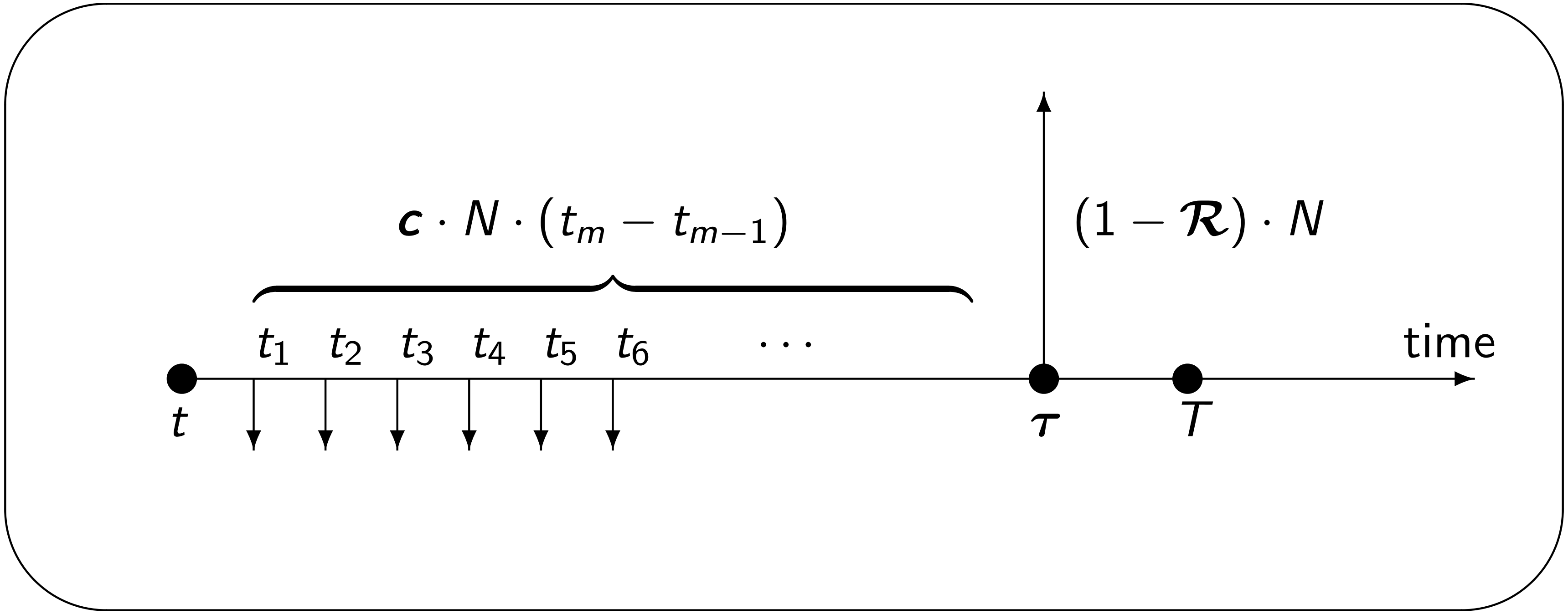


Figure: Cash flows of a single-name credit default swap

Credit default swap

Example

We consider a credit default swap, whose notional principal is \$10 mn, maturity is 5 years and payment frequency is quarterly. The credit event is the bankruptcy of the corporate entity A. We assume that the recovery rate is set to 40% and the coupon rate is equal to 2%

- 20 fixing dates: 3M, 6M, 9M, 1Y, ..., 5Y
- Quarterly premium $= \$10\,\mathrm{mn} \times 2\% \times 0.25 = \$50\,000$
- No default $\Rightarrow$ the protection buyer will pay a total of $\$50\,000 \times 20 = \$1$ mn
- The corporate defaults two years and four months after the CDS inception date $\Rightarrow$ the protection buyer will pay $9 \times \$50\,000 = \$450\,000$ and the protection seller will pay the protection leg $(1 - 40\%) \times \$10\,\mathrm{mn} = \$6$ mn

Credit default swap

If we assume that the premium is not paid after the default time $\boldsymbol{\tau}$, the stochastic discounted value of the premium leg is:

$$SV_t(\mathcal{PL}) = \sum_{t_m \geq t} \boldsymbol{c} \cdot N \cdot (t_m - t_{m-1}) \cdot \mathbb{1}\{\boldsymbol{\tau} > t_m\} \cdot e^{-\int_t^{t_m} r_s \, \mathrm{d}s}$$

The present value of the premium leg is then:

$$\begin{aligned} PV_t(\mathcal{PL}) &= \mathbb{E}\left[\left.\sum_{t_m \geq t} \boldsymbol{c} \cdot N \cdot \Delta t_m \cdot \mathbb{1}\{\boldsymbol{\tau} > t_m\} \cdot e^{-\int_t^{t_m} r_s \, \mathrm{d}s}\right| \mathcal{F}_t\right] \\ &= \sum_{t_m \geq t} \boldsymbol{c} \cdot N \cdot \Delta t_m \cdot \mathbb{E}\left[\mathbb{1}\{\boldsymbol{\tau} > t_m\}\right] \cdot \mathbb{E}\left[e^{-\int_t^{t_m} r_s \, \mathrm{d}s}\right] \\ &= \boldsymbol{c} \cdot N \cdot \sum_{t_m \geq t} \Delta t_m \mathbf{S}_t(t_m) B_t(t_m) \end{aligned}$$

where $\mathbf{S}_t(u)$ is the survival function at time u

Credit default swap

If we assume that the default leg is exactly paid at the default time $\boldsymbol{\tau}$, the stochastic discount value of the default (or protection) leg is:

$$SV_t\left(\mathcal{DL}\right) = \left(1 - \boldsymbol{\mathcal{R}}\right) \cdot N \cdot \mathbb{1}\left\{\boldsymbol{\tau} \leq T\right\} \cdot e^{-\int_t^{\boldsymbol{\tau}} r(s)\,\mathrm{d}s}$$

It follows that its present value is:

$$\begin{aligned} PV_t\left(\mathcal{DL}\right) &= \mathbb{E}\left[\left(1 - \boldsymbol{\mathcal{R}}\right) \cdot N \cdot \mathbb{1}\left\{\boldsymbol{\tau} \leq T\right\} \cdot e^{-\int_t^{\boldsymbol{\tau}} r_s\,\mathrm{d}s} \,\middle|\, \mathcal{F}_t\right] \\ &= \left(1 - \boldsymbol{\mathcal{R}}\right) \cdot N \cdot \mathbb{E}\left[\mathbb{1}\left\{\boldsymbol{\tau} \leq T\right\} \cdot B_t\left(\boldsymbol{\tau}\right)\right] \\ &= \left(1 - \boldsymbol{\mathcal{R}}\right) \cdot N \cdot \int_t^T B_t\left(u\right) f_t\left(u\right)\,\mathrm{d}u \end{aligned}$$

where $f_t\left(u\right)$ is the density function associated to the survival function $\mathbf{S}_t\left(u\right)$

Credit default swap

We deduce that the mark-to-market of the swap is:

$$\begin{aligned} P_t(T) &= PV_t(\mathcal{DL}) - PV_t(\mathcal{PL}) \\ &= (1-\mathcal{R})\, N \int_t^T B_t(u)\, f_t(u)\, \mathrm{d}u - \boldsymbol{c} N \sum_{t_m \geq t} \Delta t_m \mathbf{S}_t(t_m)\, B_t(t_m) \\ &= N \left((1-\mathcal{R}) \int_t^T B_t(u)\, f_t(u)\, \mathrm{d}u - \boldsymbol{c} \cdot \mathrm{RPV}_{01} \right) \end{aligned}$$

where $\mathrm{RPV}_{01} = \sum_{t_m \geq t} \Delta t_m \mathbf{S}_t(t_m)\, B_t(t_m)$ is called the risky PV01 and corresponds to the present value of 1 bp paid on the premium leg

CDS spread

The CDS spread $\mathcal{s}$ is the fair value coupon rate $\boldsymbol{c}$ in such a way that the initial value of the credit default swap is equal to zero $P_t = 0$:

$$\mathcal{s} = \frac{(1-\mathcal{R}) \int_t^T B_t(u)\, f_t(u)\, \mathrm{d}u}{\sum_{t_m \geq t} \Delta t_m \mathbf{S}_t(t_m)\, B_t(t_m)}$$

Credit default swap

Three properties:

1. No default risk: $\mathbf{S}_t(u) = 1 \Rightarrow \mathcal{s} = 0$
2. Recovery rate is set to 100%: $\mathcal{R} = 1 \Rightarrow \mathcal{s} = 0$
3. $\mathcal{s}$ is a decreasing function of $\mathcal{R}$

If the premium leg is paid continuously, we obtain:

$$\mathcal{s} = \frac{(1-\mathcal{R})\int_t^T B_t(u)\, f_t(u)\,\mathrm{d}u}{\int_t^T B_t(u)\,\mathbf{S}_t(u)\,\mathrm{d}u}$$

Credit default swap

If the interest rates are equal to zero ($B_t(u) = 1$) and the default times is exponential with parameter λ – $\mathbf{S}_t(u) = e^{-\lambda(u-t)}$ and $f_t(u) = \lambda e^{-\lambda(u-t)}$, we get:

$$\mathcal{s} = \frac{(1-\mathcal{R}) \cdot \lambda \cdot \int_t^T e^{-\lambda(u-t)} \,\mathrm{d}u}{\int_t^T e^{-\lambda(u-t)} \,\mathrm{d}u} = (1-\mathcal{R}) \cdot \lambda$$

If λ is relatively small, the one-year default probability is equal to:

$$\mathrm{PD} = \Pr\{\boldsymbol{\tau} \leq t+1 \mid \boldsymbol{\tau} \leq t\} = 1 - \mathbf{S}_t(t+1) = 1 - e^{-\lambda} \simeq \lambda$$

Credit triangle relationship

$$\mathcal{s} \approx (1-\mathcal{R}) \cdot \mathrm{PD}$$

$\Rightarrow$ The spread is a decreasing function of the default probability

Credit default swap

- The first CDS was traded by J.P. Morgan in 1994
- Standardization: 2003 and 2014 ISDA
- Settlement: physical or cash

In the case of physical settlement, the protection buyer delivers a bond to the protection seller and receives the notional principal amount $\Rightarrow$ the price of the defaulted bond is equal to $\mathcal{R} \cdot N \Rightarrow$ the implied mark-to-market of the physical settlement is $N - \mathcal{R} \cdot N = (1 - \mathcal{R}) \cdot N$

Credit default swap

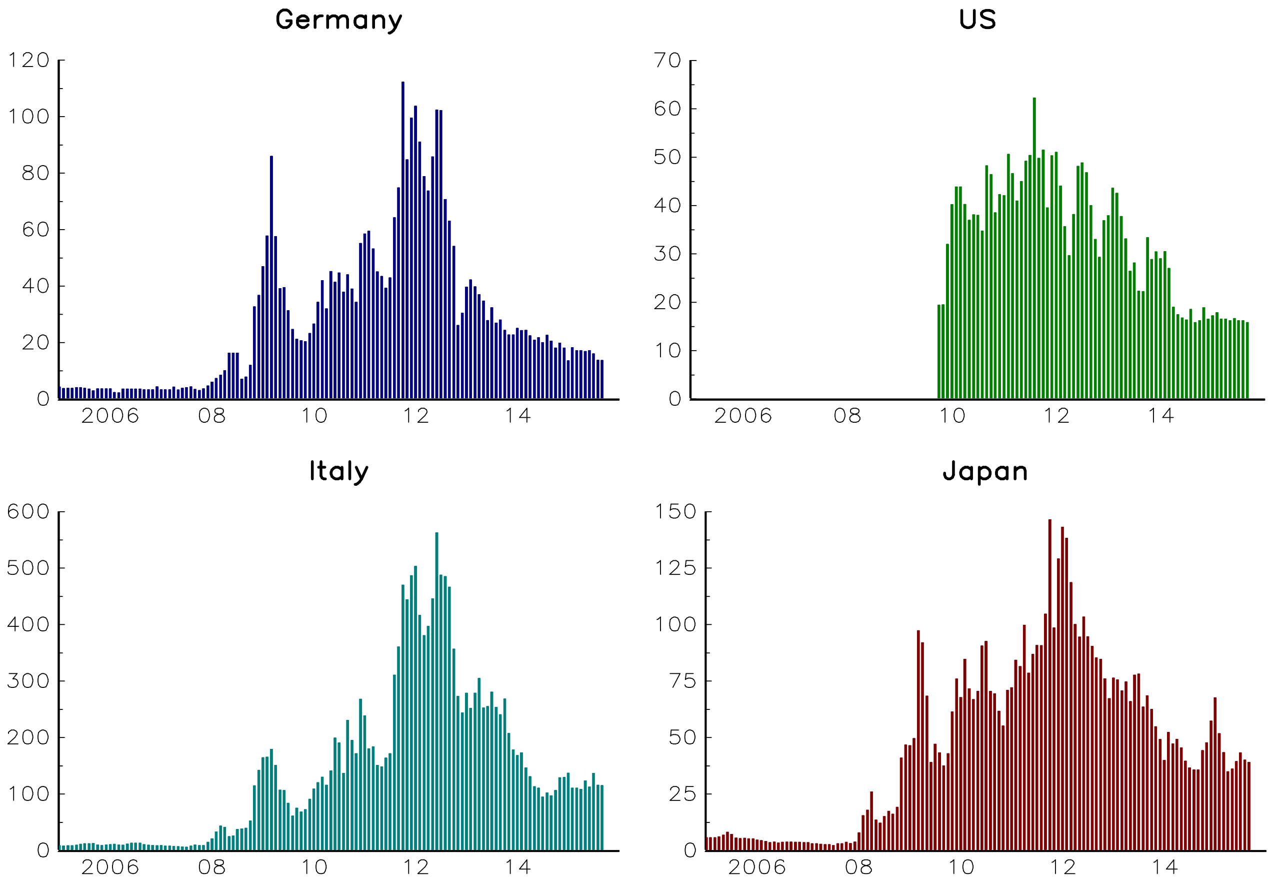


Figure: Evolution of some sovereign CDS spreads

Credit default swap

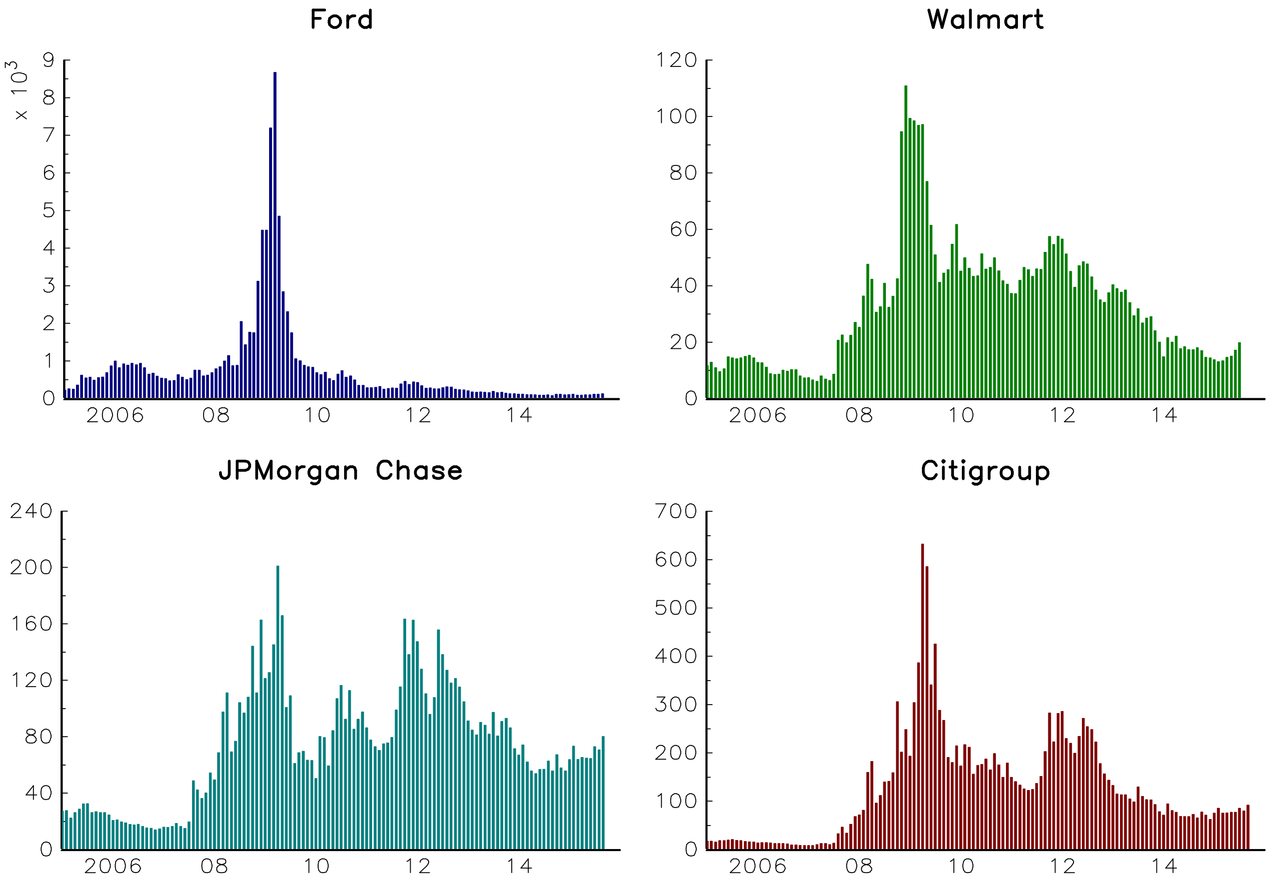


Figure: Evolution of some financial and corporate CDS spreads

Credit curve

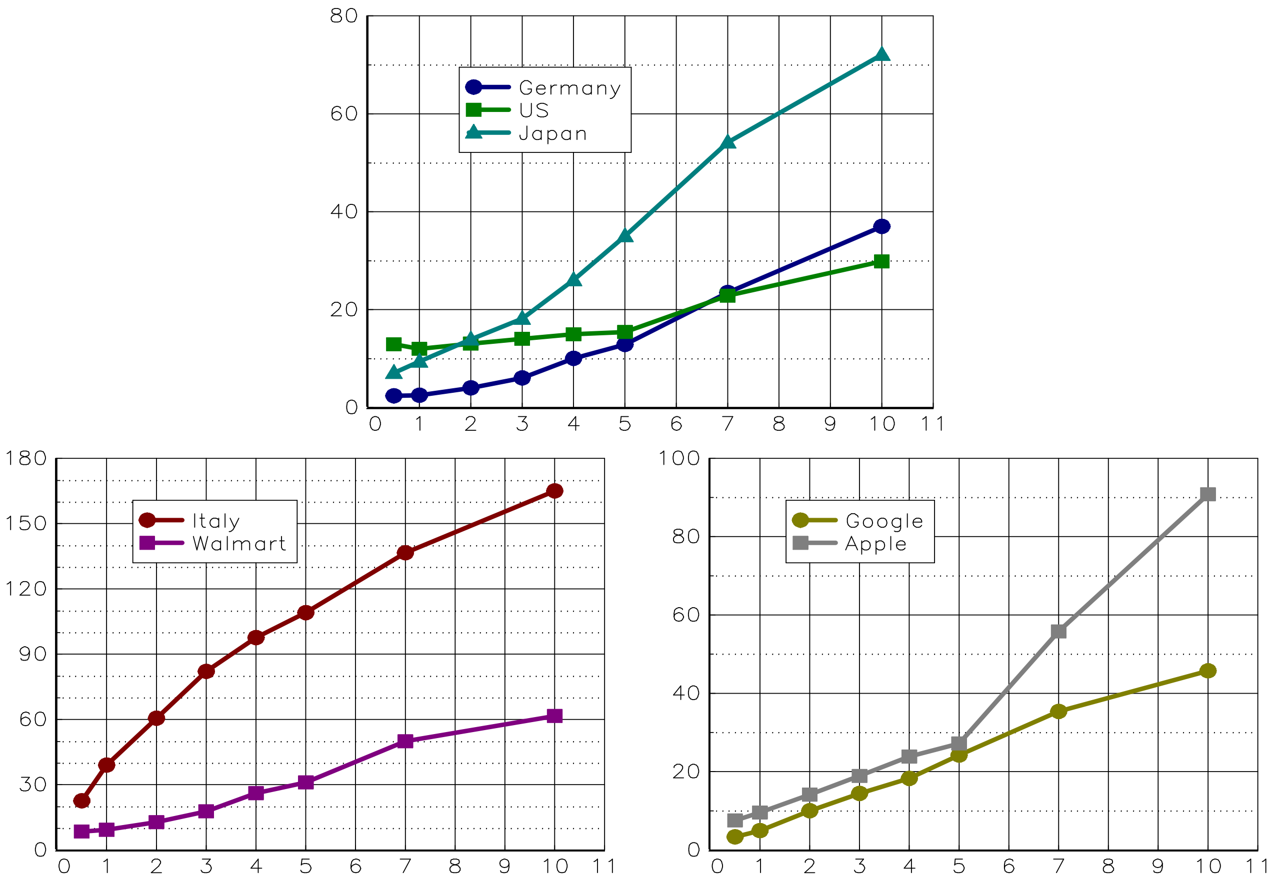


Figure: Example of CDS spread curves as of 17 September 2015

Credit risk hedging with a CDS contract

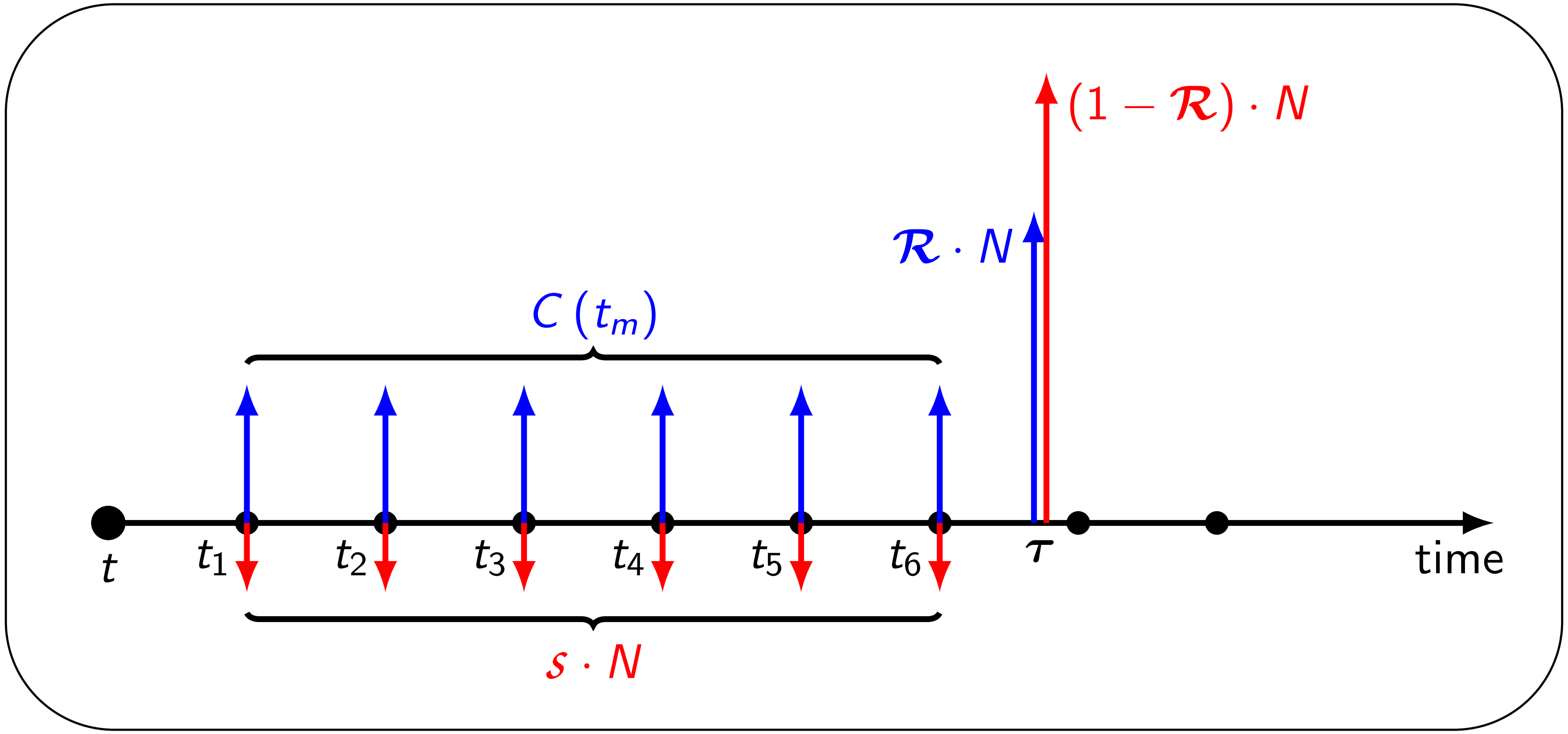


Figure: Hedging a defaultable bond with a credit default swap

$$y^{\star} = y - s \Rightarrow \textbf{CDS spread} = \textbf{Credit spread}$$

Credit risk trading with a CDS contract

Two directional trading strategies:

- '*long credit*' refers to the position of the protection seller who is exposed to the credit risk
- '*short credit*' is the position of the protection buyer who sold the credit risk of the reference entity

$\Rightarrow$ A long exposure implies that the default results in a loss, whereas a short exposure implies that the default results in a gain

Credit risk trading with a CDS contract

Let $P_{t,t'}(T)$ be the mark-to-market of a CDS position whose inception date is t, valuation date is t' and maturity date is T. We have:

$$P_{t,t}^{\mathrm{seller}}(T) = P_{t,t}^{\mathrm{buyer}}(T) = 0$$

At date $t' > t$, the mark-to-market price of the CDS is:

$$P_{t,t'}^{\mathrm{buyer}}(T) = N\left((1-\mathcal{R})\int_{t'}^{T} B_{t'}(u)\, f_{t'}(u)\,\mathrm{d}u - \mathcal{s}_t(T)\cdot \mathrm{RPV}_{01}\right)$$

whereas the value of the CDS spread satisfies the following relationship:

$$P_{t',t'}^{\mathrm{buyer}}(T) = N\left((1-\mathcal{R})\int_{t'}^{T} B_{t'}(u)\, f_{t'}(u)\,\mathrm{d}u - \mathcal{s}_{t'}(T)\cdot \mathrm{RPV}_{01}\right) = 0$$

We deduce that the P&L of the protection buyer is:

$$\Pi^{\mathrm{buyer}} = P_{t,t'}^{\mathrm{buyer}}(T) - P_{t,t}^{\mathrm{buyer}}(T) = P_{t,t'}^{\mathrm{buyer}}(T)$$

Credit risk trading with a CDS contract

We know that $P_{t',t'}^{\mathrm{buyer}}(T) = 0$ and we obtain:

$$\begin{aligned}
\Pi^{\mathrm{buyer}} &= P_{t,t'}^{\mathrm{buyer}}(T) - P_{t',t'}^{\mathrm{buyer}}(T) \\
&= N\left((1-\mathcal{R})\int_{t'}^{T} B_{t'}(u) f_{t'}(u)\,\mathrm{d}u - \mathcal{s}_t(T)\cdot \mathrm{RPV}_{01}\right) - \\
&\quad N\left((1-\mathcal{R})\int_{t'}^{T} B_{t'}(u) f_{t'}(u)\,\mathrm{d}u - \mathcal{s}_{t'}(T)\cdot \mathrm{RPV}_{01}\right) \\
&= N\cdot\left(\mathcal{s}_{t'}(T) - \mathcal{s}_t(T)\right)\cdot \mathrm{RPV}_{01}
\end{aligned}$$

Because $\Pi^{\mathrm{seller}} = -\Pi^{\mathrm{buyer}}$, we distinguish two cases:

- If $\mathcal{s}_{t'}(T) > \mathcal{s}_t(T)$, the protection buyer makes a profit, because this short credit exposure has benefited from the increase of the default risk.
- If $\mathcal{s}_{t'}(T) < \mathcal{s}_t(T)$, the protection seller makes a profit, because the default risk of the reference entity has decreased.

Credit risk trading with a CDS contract

Suppose that we are in the first case. To realize its P&L, the protection buyer has three options:

1. He could unwind the CDS exposure with the protection seller if the latter agrees. This implies that the protection seller pays the mark-to-market $P_{t,t'}^{\mathrm{buyer}}(T)$ to the protection buyer
2. He could hedge the mark-to-market value by selling a CDS on the same reference entity and the same maturity. In this situation, he continues to pay the spread $\mathcal{s}_t(T)$, but he now receives a premium, whose spread is equal to $\mathcal{s}_{t'}(T)$
3. He could reassign the CDS contract to another counterparty. The new counterparty (the protection buyer C in our case) will then pay the coupon rate $\mathcal{s}_t(T)$ to the protection seller. However, the spread is $\mathcal{s}_{t'}(T)$ at time t', which is higher than $\mathcal{s}_t(T)$. This is why the new counterparty also pays the mark-to-market $P_{t,t'}^{\mathrm{buyer}}(T)$ to the initial protection buyer

Credit risk trading with a CDS contract

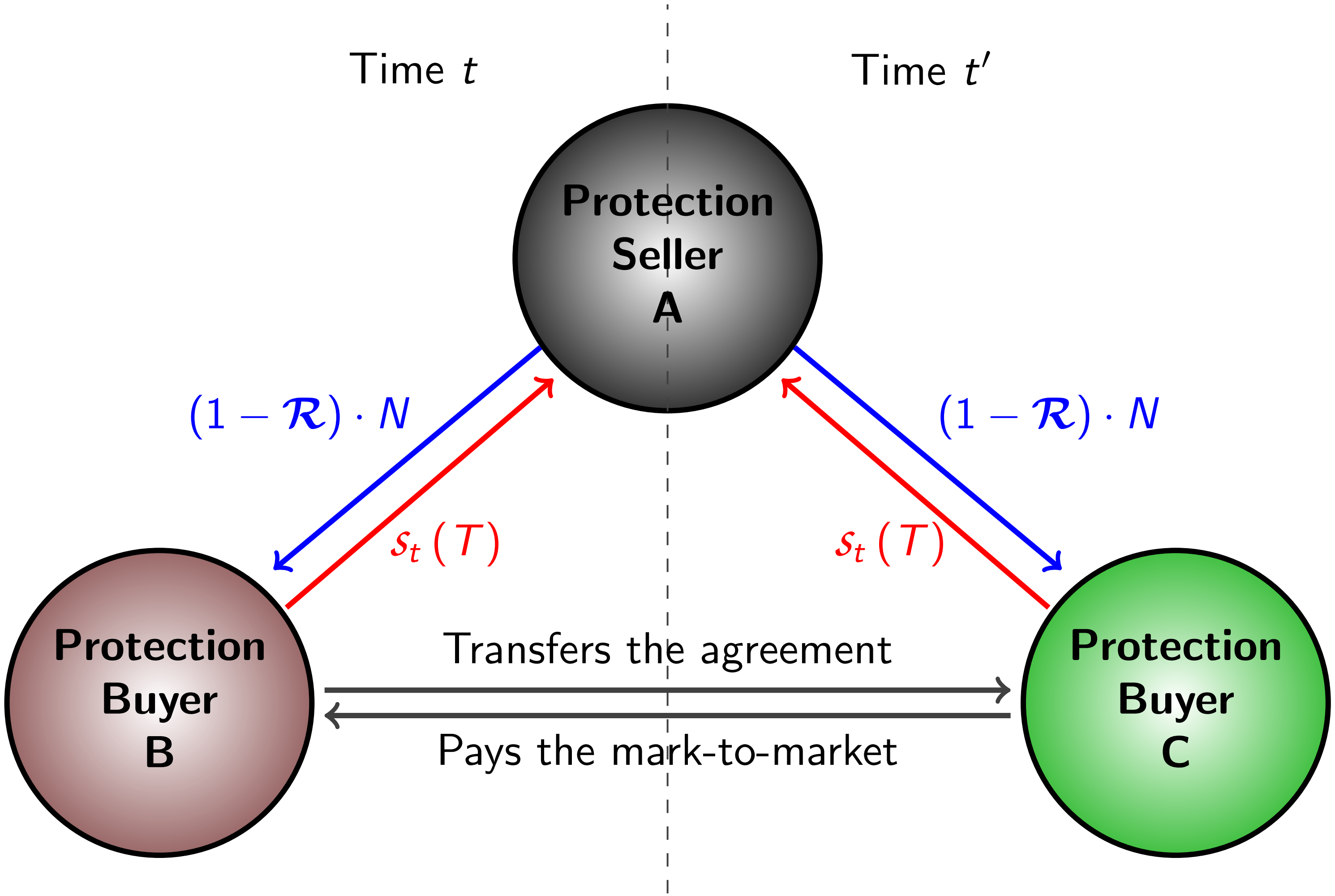


Figure: An example of CDS offsetting

Credit default swap

Example

The coupons are quarterly and the notional is equal to \$1 mn. The recovery rate $\mathcal{R}$ is set to 40% whereas the default time $\boldsymbol{\tau}$ is an exponential random variable, whose parameter λ is equal to 50 bps. We consider seven maturities (6M, 1Y, 2Y, 3Y, 5Y, 7Y and 10Y) and two coupon rates (10 and 100 bps).

Table: Price, spread and risky PV01 of CDS contracts

T	$P_t(T)$ $\boldsymbol{c}=10$	$P_t(T)$ $\boldsymbol{c}=100$	$\mathcal{s}$	RPV_{01}
1/2	998	−3 492	30.01	0.499
1	1 992	−6 963	30.02	0.995
2	3 956	−13 811	30.04	1.974
3	5 874	−20 488	30.05	2.929
5	9 527	−33 173	30.08	4.744
7	12 884	−44 804	30.10	6.410
10	17 314	−60 121	30.12	8.604

Basket default swap

- First-to-default (FtD)
- Second-to-default (StD)
- k^{th}-to-default credit derivatives

⇒ Impact of the default correlation:

$$\max\left(s_1^{\mathrm{CDS}}, \ldots, s_n^{\mathrm{CDS}}\right) \leq s^{\mathrm{FtD}} \leq \sum_{i=1}^{n} s_i^{\mathrm{CDS}}$$

Credit default indices

Definition

A credit default index is a CDS on a basket of reference entities

Table: Historical spread of CDX/iTraxx indices (in bps)

Date	CDX			iTraxx		
	NA.IG	NA.HY	EM	Europe	Japan	Asia
Dec. 2012	94.1	484.4	208.6	117.0	159.1	108.8
Dec. 2013	62.3	305.6	272.4	70.1	67.5	129.0
Dec. 2014	66.3	357.2	341.0	62.8	67.0	106.0
Sep. 2015	93.6	505.3	381.2	90.6	82.2	160.5

Credit default indices

Table: List of Markit CDX main indices

Index name	Description	n	$\mathcal{R}$
CDX.NA.IG	Investment grade entities	125	40%
CDX.NA.IG.HVOL	High volatility IG entities	30	40%
CDX.NA.XO	Crossover entities	35	40%
CDX.NA.HY	High yield entities	100	30%
CDX.NA.HY.BB	High yield BB entities	37	30%
CDX.NA.HY.B	High yield B entities	46	30%
CDX.EM	EM sovereign issuers	14	25%
LCDX	Secured senior loans	100	70%
MCDX	Municipal bonds	50	80%

Credit default indices

Table: List of Markit iTraxx main indices

Index name	Description	n	$\mathcal{R}$
iTraxx Europe	European IG entities	125	40%
iTraxx Europe HiVol	European HVOL IG entities	30	40%
iTraxx Europe Crossover	European XO entities	40	40%
iTraxx Asia	Asian (ex-Japan) IG entities	50	40%
iTraxx Asia HY	Asian (ex-Japan) HY entities	20	25%
iTraxx Australia	Australian IG entities	25	40%
iTraxx Japan	Japanese IG entities	50	35%
iTraxx SovX G7	G7 governments	7	40%
iTraxx LevX	European leveraged loans	40	40%

Collateralized debt obligation (CDO)

A CDO is a pay-through ABS structure, whose securities are bonds linked to a series of tranches

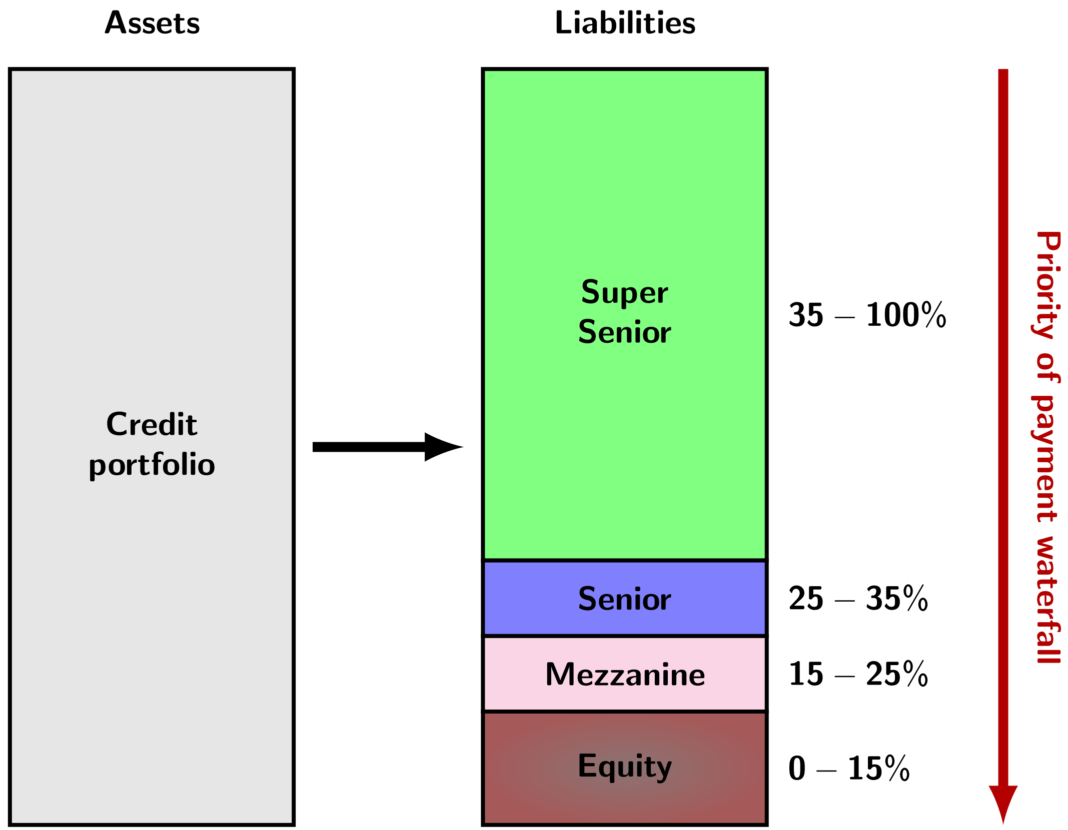


Figure: An example of a CDO structure

Collateralized debt obligation (CDO)

The returns of the 4 bonds depend on the loss of the corresponding tranche. Each tranche is characterized by an attachment point A and a detachment point D. In our example, we have:

Tranche	Equity	Mezzanine	Senior	Super senior
A	0%	15%	25%	35%
D	15%	25%	35%	100%

The protection buyer of the tranche $[A, D]$ pays a coupon rate $\mathbf{c}^{[A,D]}$ on the nominal outstanding amount of the tranche to the protection seller. In return, he receives the protection leg, which is the loss of the tranche $[A, D]$

CDO pricing

We have:

$$L_t(u) = \sum_{i=1}^{n} N_i \cdot (1 - \boldsymbol{\mathcal{R}}_i) \cdot \mathbb{1}\{\boldsymbol{\tau}_i \leq u\}$$

and:

$$L_t^{[A,D]}(u) = (L_t(u) - A) \cdot \mathbb{1}\{A \leq L_t(u) \leq D\} + (D - A) \cdot \mathbb{1}\{L_t(u) > D\}$$

The nominal outstanding amount of the tranche is therefore:

$$N_t^{[A,D]}(u) = (D - A) - L_t^{[A,D]}(u)$$

The spread of the CDO tranche is

$$\mathcal{s}^{[A,D]} = \frac{\mathbb{E}\left[\sum_{t_m \geq t} \Delta L_t^{[A,D]}(t_m) \cdot B_t(t_m)\right]}{\mathbb{E}\left[\sum_{t_m \geq t} \Delta t_m \cdot N_t^{[A,D]}(t_m) \cdot B_t(t_m)\right]}$$

We obviously have the following inequalities

$$\mathcal{s}^{\text{Equity}} > \mathcal{s}^{\text{Mezzanine}} > \mathcal{s}^{\text{Senior}} > \mathcal{s}^{\text{Super senior}}$$

Credit risk

It is the risk of loss on a debt instrument resulting from the failure of the borrower to make required payments: *default risk* $\neq$ *downgrading risk*

Definition (BCBS, 2006)

A default is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place

- The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held)
- The obligor is past due more than 90 days on any material credit obligation to the banking group. Overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstandings

A fair game?

Table: World's largest banks in 1981 and 1988

	1981		1988	
	Bank	Assets	Bank	Assets
1	Bank of America (US)	115.6	Dai-Ichi Kangyo (JP)	352.5
2	Citicorp (US)	112.7	Sumitomo (JP)	334.7
3	BNP (FR)	106.7	Fuji (JP)	327.8
4	Crédit Agricole (FR)	97.8	Mitsubishi (JP)	317.8
5	Crédit Lyonnais (FR)	93.7	Sanwa (JP)	307.4
6	Barclays (UK)	93.0	Industrial Bank (JP)	261.5
7	Société Générale (FR)	87.0	Norinchukin (JP)	231.7
8	Dai-Ichi Kangyo (JP)	85.5	Crédit Agricole (FR)	214.4
9	Deutsche Bank (DE)	84.5	Tokai (JP)	213.5
10	National Westminster (UK)	82.6	Mitsubishi Trust (JP)	206.0

The Basel I framework

Table: Risk weights by category of on-balance sheet assets

RW	Instruments
0%	Cash Claims on central governments and central banks denominated in national currency and funded in that currency Other claims on OECD central governments and central banks Claims† collateralized by cash of OECD government securities
20%	Claims† on multilateral development banks Claims† on banks incorporated in the OECD and claims guaranteed by OECD incorporated banks Claims† on securities firms incorporated in the OECD subject to comparable supervisory and regulatory arrangements Claims† on banks incorporated in countries outside the OECD with a residual maturity of up to one year Claims† on non-domestic OECD public-sector entities Cash items in process of collection
50%	Loans fully secured by mortgage on residential property
100%	Claims on the private sector Claims on banks incorporated outside the OECD with a residual maturity of over one year Claims on central governments outside the OECD and non denominated in national currency All other assets

The Basel I framework

For off-balance sheet assets, the amount E of a credit line is converted to an exposure at default:

$$\mathrm{EAD} = E \cdot \mathrm{CCF}$$

where CCF is the credit conversion factor (100%, 50%, 20% and 0%)

The Basel I framework

Table: Illustration of capital requirement

Balance Sheet	Asset	E	CCF	EAD	RW	RWA
On-	US bonds			100	0%	0
	Mexico bonds			20	100%	20
	Argentine debt			20	0%	0
	Home mortgage			500	50%	250
	Corporate loans			500	100%	500
	Credit lines			40	100%	40
Off-	Standby facilities	20	100%	20	0%	0
	Credit lines ($>$ 1Y)	42	50%	21	100%	21
	Credit lines ($\leq$ 1Y)	18	0%	0	100%	0
	Total					831

The Basel II framework

- The standardized approach (SA)
- The internal ratings-based approach (IRB)

The Basel II standardized approach

Table: Risk weights of the SA approach (Basel II)

Rating		AAA to AA−	A+ to A−	BBB+ to BBB−	BB+ to B−	CCC+ to C	NR
Sovereigns		0%	20%	50%	100%	150%	100%
Banks	1	20%	50%	100%	100%	150%	100%
	2	20%	50%	50%	100%	150%	50%
	2 ST	20%	20%	20%	50%	150%	20%
Corporates				BBB+ to BB−		B+ to C	
		20%	50%	100%		150%	100%
Retail				75%			
Residential mortgages				35%			
Commercial mortgages				100%			

The Basel II standardized approach

Table: Comparison of risk weights between Basel I and Basel II

Entity	Rating	Maturity	Basel I	Basel II
Sovereign (OECD)	AAA		0%	0%
Sovereign (OECD)	A-		0%	20%
Sovereign	BBB		100%	50%
Bank (OECD)	BBB	2Y	20%	50%
Bank	BBB	2M	100%	20%
Corporate	AA+		100%	20%
Corporate	BBB		100%	100%

Credit ratings

Table: Credit rating system of S&P, Moody's and Fitch

	Prime Maximum Safety	High Grade High Quality			Upper Medium Grade		
S&P/Fitch	AAA	AA+	AA	AA−	A+	A	A−
Moody's	Aaa	Aa1	Aa2	Aa3	A1	A2	A3

	Lower Medium Grade			Non Investment Grade Speculative		
S&P/Fitch	BBB+	BBB	BBB−	BB+	BB	BB−
Moody's	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3

	Highly Speculative			Substantial Risk	In Poor Standing		Extremely Speculative
S&P/Fitch	B+	B	B−	CCC+	CCC	CCC−	CC
Moody's	B1	B2	B3	Caa1	Caa2	Caa3	Ca

Credit ratings

Table: Examples of country's S&P rating

Country	Local currency		Foreign currency	
	Jun. 2009	Oct. 2015	Jun. 2009	Oct. 2015
Argentina	B-	CCC+	B-	SD
Brazil	BBB+	BBB-	BBB-	BB+
China	A+	AA-	A+	AA-
France	AAA	AA	AAA	AA
Italy	A+	BBB-	A+	BBB-
Japan	AA	A+	AA	A+
Russia	BBB+	BBB-	BBB	BB+
Spain	AA+	BBB+	AA+	BBB+
Ukraine	B-	CCC+	CCC+	SD
US	AAA	AA+	AA+	AA+

The Basel II standardized approach

CCF (Basel II $\approx$ Basel I)

Credit risk mitigation

1. Collateralized transactions
2. Guarantees and credit derivatives

Credit risk mitigation

Collateralized transactions

1. Cash and comparable instruments
2. Gold
3. Debt securities which are rated AAA to BB- when issued by sovereigns or AAA to BBB- when issued by other entities or at least A-3/P-3 for short-term debt instruments
4. Debt securities which are not rated but fulfill certain criteria (senior debt issued by banks, listed on a recognisee exchange and sufficiently liquid)
5. Equities that are included in a main index
6. UCITS and mutual funds, whose assets are eligible instruments and which offer a daily liquidity
7. Equities which are listed on a recognized exchange and UCITS/mutual funds which include such equities

Credit risk mitigation

Collateralized transactions

Simple approach

$$\mathrm{RWA} = (\mathrm{EAD} - C) \cdot \mathrm{RW} + C \cdot \max(\mathrm{RW}_C, 20\%)$$

where EAD is the exposure at default, C is the market value of the collateral, RW is the risk weight appropriate to the exposure and RW_C is the risk weight of the collateral

Comprehensive approach

The risk-weighted asset amount after risk mitigation is $\mathrm{RWA} = \mathrm{RW} \cdot \mathrm{EAD}^{\star}$ whereas $\mathrm{EAD}^{\star}$ is the modified exposure at default:

$$\mathrm{EAD}^{\star} = \max(0, (1 + H_E) \cdot \mathrm{EAD} - (1 - H_C - H_{FX}) \cdot C)$$

where H_E is the haircut applied to the exposure, H_C is the haircut applied to the collateral and H_{FX} is the haircut for currency risk

Credit risk mitigation

Collateralized transactions

Table: Standardized supervisory haircuts for collateralized transactions

Rating	Residual Maturity	Sovereigns	Others
AAA to AA−	0−1Y	0.5%	1%
	1−5Y	2%	4%
	5Y+	4%	8%
A+ to BBB−	0−1Y	1%	2%
	1−5Y	3%	6%
	5Y+	6%	12%
BB+ to BB−		15%	
Cash		0%	
Gold		15%	
Main index equities		15%	
Equities listed on a recognized exchange		25%	
FX risk		8%	

Credit risk mitigation

Guarantees and credit derivatives

Banks can use these credit protection instruments if they are direct, explicit, irrevocable and unconditional

Simple approach

$$\mathrm{RWA} = (\mathrm{EAD} - C) \cdot \mathrm{RW} + C \cdot \max\left(\mathrm{RW}_C, 20\%\right)$$

where EAD is the exposure at default, C is the market value of the collateral, RW is the risk weight appropriate to the exposure and RW_C is the risk weight of the collateral

The Basel II internal ratings-based approach

4 parameters:

- the exposure at default (EAD)
- the probability of default (PD)
- the loss given default (LGD)
- the effective maturity (M)

The credit risk measure is the sum of individual risk contributions:

$$\mathcal{R}(w) = \sum_{i=1}^{n} \mathcal{RC}_i$$

where $\mathcal{RC}_i$ is a function of the four risk components:

$$\mathcal{RC}_i = f_{\mathrm{IRB}}(\mathrm{EAD}_i, \mathrm{LGD}_i, \mathrm{PD}_i, \mathrm{M}_i)$$

and f_{IRB} is the IRB fomula

IRB is not an internal model, but an external model with internal parameters

The Basel II internal ratings-based approach

The mechanism of the IRB approach is the following:

- a classification of exposures (sovereigns, banks, corporates, retail portfolios, etc.)
- for each credit i, the bank estimates the probability of default
- it uses the standard regulatory values of the other risk components (EAD_i, LGD_i and M_i) or estimates them in the case of AIRB
- the bank calculate then the risk-weighted assets RWA_i of the credit by applying the right IRB formula f_{IRB} to the risk components

$\Rightarrow$ Distinction between FIRB (foundation IRB) and AIRB (advanced IRB)

$\Rightarrow$ **Internal ratings are central to the IRB approach**

The Basel II internal ratings-based approach

Table: An example of internal rating system

Rating		Degree of risk	Definition	Borrower category by self-assessment
1		No essential risk	Extremely high degree of certainty of repayment	Normal
2		Negligible risk	High degree of certainty of repayment	
3		Some risk	Sufficient certainty of repayment	
4	A B	Better than average	There is certainty of repayment but substantial changes in the environment in the future may have some impact on this uncertainty	
5	A B	Average	There are no problems foreseeable in the future, but a strong likelihood of impact from changes in the environment	
6	A B	Tolerable	There are no problems foreseeable in the future, but the future cannot be considered entirely safe	
7		Lower than average	There are no problems at the current time but the financial position of the borrower is relatively weak	
8	A B	Needs preventive management	There are problems with lending terms or fulfilment, or the borrower's business conditions are poor or unstable, or there are other factors requiring careful management	Needs attention
9		Needs serious management	There is a high likelihood of bankruptcy in the future	In danger of bankruptcy
10	I		The borrower is in serious financial straits and "effectively bankrupt"	Effectively bankruptcy
	II		The borrower is bankrupt	Bankrupt

The Basel II internal ratings-based approach

Another example of internal rating system

The rating system of Crédit Agricole is:

- A+, A,
- B+, B,
- C+, C, C-,
- D+, D, D-,
- E+, E and E-

Source: Crédit Agricole, Annual Financial Report 2014, page 201

The credit risk model of Basel II

Assumptions

- The portfolio loss is equal to:

$$L = \sum_{i=1}^{n} w_i \cdot \mathrm{LGD}_i \cdot \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

  where w_i and T_i are the exposure at default and the residual maturity of the i^{th} credit
- The loss given default LGD_i is a random variable
- The default time $\boldsymbol{\tau}_i$ depends on a set of risk factors X, whose probability distribution is denoted by $\mathbf{H}$
- Let $p_i(X)$ be the conditional default probability. The (unconditional or long-term) default probability is:

$$p_i = \mathbb{E}_X\left[\mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}\right] = \mathbb{E}_X\left[p_i(X)\right]$$

- Let $D_i = \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$ be the default indicator function. Conditionally to the risk factors X, D_i is a Bernoulli random variable with probability $p_i(X)$

The credit risk model of Basel II

Under the standard assumptions that the loss given default is independent from the default time and the default times are conditionally independent, we obtain:

$$\mathbb{E}\left[L \mid X\right] = \sum_{i=1}^{n} w_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot \mathbb{E}\left[D_i \mid X\right] = \sum_{i=1}^{n} w_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot p_i\left(X\right)$$

The credit risk model of Basel II

We also have (HFRM, Exercise 3.4.8, page 255):

$$\sigma^2 (L \mid X) = \sum_{i=1}^{n} w_i^2 \cdot \left(\mathbb{E}\left[\mathrm{LGD}_i^2 \right] \cdot \mathbb{E}\left[D_i^2 \mid X \right] - \mathbb{E}^2 \left[\mathrm{LGD}_i \right] \cdot p_i^2 (X) \right)$$

Since we have:

$$\begin{array}{rcl} \mathbb{E}\left[D_i^2 \mid X \right] & = & p_i (X) \\ \mathbb{E}\left[\mathrm{LGD}_i^2 \right] & = & \sigma^2 (\mathrm{LGD}_i) + \mathbb{E}^2 \left[\mathrm{LGD}_i \right] \end{array}$$

we deduce that:

$$\sigma^2 (L \mid X) = \sum_{i=1}^{n} w_i^2 \cdot A_i$$

where:

$$A_i = \mathbb{E}^2 \left[\mathrm{LGD}_i \right] \cdot p_i (X) \cdot (1 - p_i (X)) + \sigma^2 (\mathrm{LGD}_i) \cdot p_i (X)$$

The credit risk model of Basel II

The concept of granularity

Infinitely granular portfolio

The portfolio is infinitely fine-grained if there is no concentration risk:

$$\lim_{n \to \infty} \max \frac{w_i}{\sum_{j=1}^{n} w_j} = 0$$

$\Rightarrow$ the conditional distribution of L degenerates to its conditional expectation $\mathbb{E}[L \mid X]$

The intuition of this result is the following: We consider a fine-grained portfolio equivalent to the original portfolio by replacing the original credit i by m credits with the same default probability p_i, the same loss given default LGD_i but an exposure at default divided by m. Let L_m be the loss of the equivalent fine-grained portfolio. When m tends to ∞, we obtain the infinitely fine-grained portfolio. Conditionally to the risk factors X, the portfolio loss L_∞ is equal to the conditional mean $\mathbb{E}[L \mid X]$

The credit risk model of Basel II

Proof

We have:

$$\mathbb{E}\left[L_m \mid X\right] = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \frac{w_i}{m} \right) \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot \mathbb{E}\left[D_i \mid X\right] = \mathbb{E}\left[L \mid X\right]$$

and:

$$\sigma^2\left(L_m \mid X\right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \frac{w_i^2}{m^2} \right) \cdot A_i = \frac{1}{m} \sum_{i=1}^{n} w_i^2 \cdot A_i = \frac{1}{m} \sigma^2\left(L_m \mid X\right)$$

We note that $\mathbb{E}\left[L_\infty \mid X\right] = \mathbb{E}\left[L \mid X\right]$ and $\sigma^2\left(L_\infty \mid X\right) = 0$. Conditionally to the risk factors X, the portfolio loss L_∞ is equal to the conditional mean $\mathbb{E}\left[L \mid X\right]$

The credit risk model of Basel II

The associated probability distribution $\mathbf{F}$ is then:

$$\begin{aligned}\mathbf{F}(\ell) &= \Pr\{L_\infty \leq \ell\} \\ &= \Pr\{\mathbb{E}[L \mid X] \leq \ell\} \\ &= \Pr\left\{\sum_{i=1}^{n} w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot p_i(X) \leq \ell\right\}\end{aligned}$$

Let $g(x)$ be the function $\sum_{i=1}^{n} w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot p_i(x)$. We have:

$$\mathbf{F}(\ell) = \int \cdots \int \mathbb{1}\{g(x) \leq \ell\} \, \mathrm{d}\mathbf{H}(x)$$

$\Rightarrow$ Not possible to obtain a closed-form formula for the value-at-risk $\mathbf{F}^{-1}(\alpha)$:

$$\mathbf{F}^{-1}(\alpha) = \{\ell : \Pr\{g(X) \leq \ell\} = \alpha\}$$

The credit risk model of Basel II

The single risk factor case

If we consider a single risk factor and assume that $g(x)$ is an increasing function, we obtain:

$$\begin{aligned} \Pr\{g(X) \leq \ell\} = \alpha & \Leftrightarrow \Pr\left\{X \leq g^{-1}(\ell)\right\} = \alpha \\ & \Leftrightarrow \mathbf{H}\left(g^{-1}(\ell)\right) = \alpha \\ & \Leftrightarrow \ell = g\left(\mathbf{H}^{-1}(\alpha)\right) \end{aligned}$$

We finally deduce that the value-at-risk has the following expression:

$$\mathbf{F}^{-1}(\alpha) = g\left(\mathbf{H}^{-1}(\alpha)\right) = \sum_{i=1}^{n} w_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot p_i\left(\mathbf{H}^{-1}(\alpha)\right)$$

The credit risk model of Basel II

Euler allocation principle

The value-at-risk satisfies the Euler allocation principle:

$$\mathbf{F}^{-1}(\alpha) = \sum_{i=1}^{n} \mathcal{RC}_i$$

where the expression of the risk contribution is:

$$\mathcal{RC}_i = w_i \cdot \frac{\partial\, \mathbf{F}^{-1}(\alpha)}{\partial\, w_i} = w_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot p_i\left(\mathbf{H}^{-1}(\alpha)\right)$$

The credit risk model of Basel II

Remark

If $g(x)$ *is a decreasing function, we obtain* $\Pr\left\{X \geq g^{-1}(\ell)\right\} = \alpha$ *and:*

$$\mathbf{F}^{-1}(\alpha) = \sum_{i=1}^{n} w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot p_i\left(\mathbf{H}^{-1}(1-\alpha)\right)$$

The risk contribution becomes:

$$\mathcal{RC}_i = w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot p_i\left(\mathbf{H}^{-1}(1-\alpha)\right)$$

The credit risk model of Basel II

Summary

Under the assumptions:

- $\mathcal{H}_1$ The loss given default LGD_i is independent from the default time $\boldsymbol{\tau}_i$
- $\mathcal{H}_2$ The default times $(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_n)$ depend on a single risk factor X and are conditionally independent with respect to X
- $\mathcal{H}_3$ The portfolio is infinitely fine-grained, meaning that there is no exposure concentration

we have:

$$\mathcal{RC}_i = w_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot p_i\left(\mathbf{H}^{-1}(\pi)\right)$$

where $\pi = \alpha$ if $p_i(X)$ is an increasing function of X or $\pi = 1 - \alpha$ if $p_i(X)$ is a decreasing function of X

The credit risk model of Basel II

Closed-form formula of the value-at-risk

$\Rightarrow$ Merton (1974) / Vasicek (1991)

Let Z_i be the normalized asset value of the entity i. In the Merton model, the default occurs when Z_i is below a given barrier B_i: $D_i = 1 \Leftrightarrow Z_i < B_i$. By assuming that Z_i is Gaussian, we deduce that:

$$p_i = \Pr\left\{D_i = 1\right\} = \Pr\left\{Z_i < B_i\right\} = \Phi\left(B_i\right)$$

and $B_i = \Phi^{-1}\left(p_i\right)$

We assume that the asset value Z_i depends on the common risk factor X and an idiosyncratic risk factor ε_i as follows:

$$Z_i = \sqrt{\rho} X + \sqrt{1-\rho}\varepsilon_i$$

X and ε_i are two independent standard normal random variables and we have:

$$\mathbb{E}\left[Z_i Z_j\right] = \mathbb{E}\left[\left(\sqrt{\rho} X + \sqrt{1-\rho}\varepsilon_i\right)\left(\sqrt{\rho} X + \sqrt{1-\rho}\varepsilon_j\right)\right] = \rho$$

where ρ is the constant asset correlation

The credit risk model of Basel II

Closed-form formula of the value-at-risk

The conditional default probability is equal to:

$$\begin{aligned} p_i(X) := \Pr\{D_i = 1 \mid X\} &= \Pr\{Z_i < B_i \mid X\} \\ &= \Pr\left\{\sqrt{\rho}X + \sqrt{1-\rho}\varepsilon_i < B_i\right\} \\ &= \Pr\left\{\varepsilon_i < \frac{B_i - \sqrt{\rho}X}{\sqrt{1-\rho}}\right\} \\ &= \Phi\left(\frac{B_i - \sqrt{\rho}X}{\sqrt{1-\rho}}\right) \end{aligned}$$

We obtain:

$$g(x) = \sum_{i=1}^{n} w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot p_i(x) = \sum_{i=1}^{n} w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right)$$

Since $g(x)$ is a decreasing function if $w_i \geq 0$, we have:

$$\mathcal{RC}_i = w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot \Phi\left(\frac{\Phi^{-1}(p_i) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right)$$

The credit risk model of Basel II

Theorem (HFRM, Appendix A.2.2.5, page 1063)

$$\int_{-\infty}^{c} \Phi(a+bx)\,\phi(x)\,\mathrm{d}x = \Phi_2\left(c, \frac{a}{\sqrt{1+b^2}}; \frac{-b}{\sqrt{1+b^2}}\right)$$

p_i is the unconditional default probability

We verify that:

$$\begin{aligned}
\mathbb{E}_X\left[p_i(X)\right] &= \mathbb{E}_X\left[\Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}X}{\sqrt{1-\rho}}\right)\right] \\
&= \int_{-\infty}^{\infty} \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right)\phi(x)\,\mathrm{d}x \\
&= \Phi_2\left(\infty, \frac{\Phi^{-1}(p_i)}{\sqrt{1-\rho}} \cdot \left(\frac{1}{1-\rho}\right)^{-1/2}; \frac{\sqrt{\rho}}{\sqrt{1-\rho}}\left(\frac{1}{1-\rho}\right)^{-1/2}\right) \\
&= \Phi_2\left(\infty, \Phi^{-1}(p_i); \sqrt{\rho}\right) = \Phi\left(\Phi^{-1}(p_i)\right) = p_i
\end{aligned}$$

The credit risk model of Basel II

Example

We consider a homogeneous portfolio with 100 credits. For each credit, the exposure at default, the expected LGD and the probability of default are set to \$1 mn, 50% and 5%

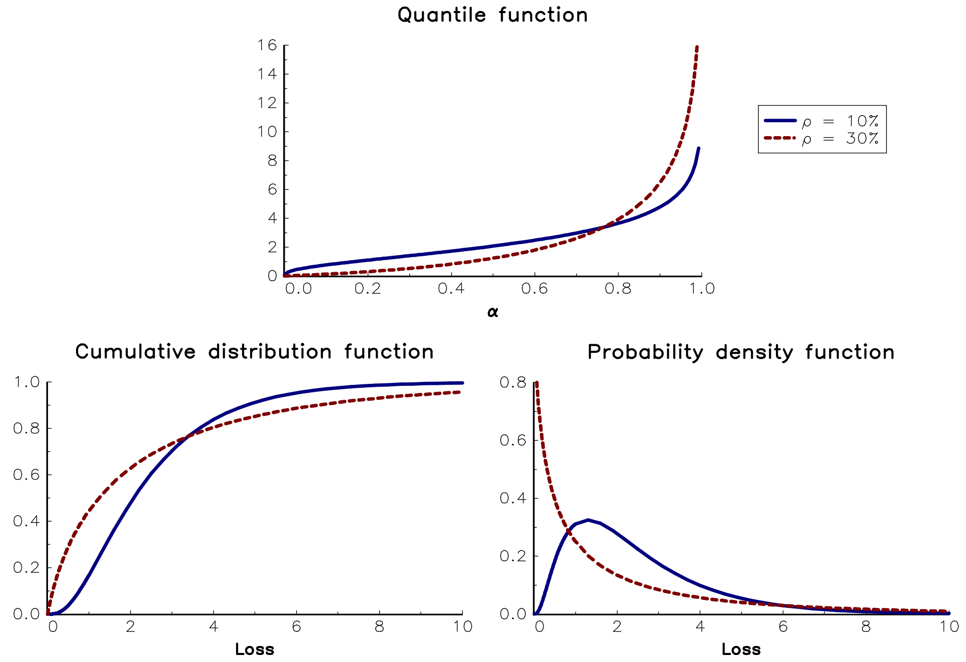


Figure: Probability functions of the credit portfolio loss

The credit risk model of Basel II

What is the impact of the maturity?

the maturity T_i is taken into account through the probability of default $\Rightarrow$ $p_i = \Pr\{\boldsymbol{\tau}_i \leq T_i\}$

Let us denote PD_i the annual default probability of the obligor. If we assume that the default time is Markovian, we have the following relationship:

$$p_i = 1 - \Pr\{\boldsymbol{\tau}_i > T_i\} = 1 - (1 - \mathrm{PD}_i)^{T_i}$$

We deduce that:

$$\mathcal{RC}_i = w_i \cdot \mathbb{E}[\mathrm{LGD}_i] \cdot \Phi\left(\frac{\Phi^{-1}\left(1 - (1 - \mathrm{PD}_i)^{T_i}\right) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right)$$

The credit risk model of Basel II

Maturity adjustment

The maturity adjustment is the function $\varphi(t)$ such that $\varphi(1) = 1$ and:

$$\mathcal{RC}_i \approx w_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot \Phi\left(\frac{\Phi^{-1}\left(\mathrm{PD}_i\right) + \sqrt{\rho}\Phi^{-1}\left(\alpha\right)}{\sqrt{1-\rho}}\right) \cdot \varphi\left(T_i\right)$$

The IRB formulas

A long process to obtain the finalized formulas

- January 2001: $\alpha = 99.5\%$, $\rho = 20\%$ and a standard maturity of three years
- April 2001: **Quantitative Impact Study** (QIS)
- November 2001: Results of the QIS 2

Table: Percentage change in capital requirements under CP2 proposals

		SA	FIRB	AIRB
G10	Group 1	6%	14%	-5%
	Group 2	1%		
EU	Group 1	6%	10%	-1%
	Group 2	-1%		
Others		5%		

- July 2002: QIS 2.5
- May 2003: QIS 3
- June 2004: Basel II

The IRB formulas

If we use the notations of the Basel Committee, the risk contribution has the following expression:

$$\mathcal{RC} = \mathrm{EAD} \cdot \mathrm{LGD} \cdot \Phi \left(\frac{\Phi^{-1}\left(1 - (1 - \mathrm{PD})^{\mathrm{M}}\right) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}} \right)$$

where:

- EAD is the exposure at default
- LGD is the (expected) loss given default
- PD is the (one-year) probability of default
- M is the effective maturity

The IRB formulas

Because $\mathcal{RC}$ is directly the capital requirement ($\mathcal{RC} = 8\% \times \mathrm{RWA}$), we deduce that the risk-weighted asset amount is equal to:

$$\mathrm{RWA} = 12.50 \cdot \mathrm{EAD} \cdot \mathcal{K}^{\star}$$

where $\mathcal{K}^{\star}$ is the normalized required capital for a unit exposure:

$$\mathcal{K}^{\star} = \mathrm{LGD} \cdot \Phi \left(\frac{\Phi^{-1}\left(1 - (1 - \mathrm{PD})^{\mathrm{M}}\right) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}} \right)$$

The IRB formulas

In order to obtain the finalized formulas, the Basel Committee has introduced the following modifications:

- A maturity adjustment $\varphi(\mathrm{M})$ has been added:

$$\mathcal{K}^{\star} \approx \mathrm{LGD} \cdot \Phi \left(\frac{\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}} \right) \cdot \varphi(\mathrm{M})$$

- The confidence level is 99.9% instead of 99.5%
- The default correlation is a parametric function $\rho(\mathrm{PD})$ in order that low ratings are not too penalizing for capital requirements;
- The credit risk measure is the unexpected loss:

$$\mathrm{UL}_{\alpha} = \mathrm{VaR}_{\alpha} - \mathbb{E}[L]$$

Final supervisory formula

$$\mathcal{K}^{\star} = \left(\mathrm{LGD} \cdot \Phi \left(\frac{\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho(\mathrm{PD})}\Phi^{-1}(99.9\%)}{\sqrt{1-\rho(\mathrm{PD})}} \right) - \mathrm{LGD} \cdot \mathrm{PD} \right) \cdot \varphi(\mathrm{M})$$

The IRB formulas

Risk-weighted assets for corporate, sovereign, and bank exposures

The three asset classes use the same formula:

$$\mathcal{K}^\star = \left(\mathrm{LGD} \cdot \Phi\left(\frac{\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho(\mathrm{PD})}\Phi^{-1}(99.9\%)}{\sqrt{1-\rho(\mathrm{PD})}}\right) - \mathrm{LGD} \cdot \mathrm{PD}\right) \cdot \left(\frac{1 + (\mathrm{M} - 2.5) \cdot b(\mathrm{PD})}{1 - 1.5 \cdot b(\mathrm{PD})}\right)$$

with:

$$b(\mathrm{PD}) = (0.11852 - 0.05478 \cdot \ln(\mathrm{PD}))^2$$

and:

$$\rho(\mathrm{PD}) = 12\% \times \left(\frac{1 - e^{-50 \times \mathrm{PD}}}{1 - e^{-50}}\right) + 24\% \times \left(1 - \frac{1 - e^{-50 \times \mathrm{PD}}}{1 - e^{-50}}\right)$$

The IRB formulas

Risk-weighted assets for small and medium-sized enterprises

SMEs are defined as corporate entities where the reported sales for the consolidated group of which the firm is a part is less than 50 € mn

⇒ New parametric function for the default correlation:

$$\rho^{\mathrm{SME}}(\mathrm{PD}) = \rho(\mathrm{PD}) - 0.04 \cdot \left(1 - \frac{(\max(S, 5) - 5)}{45}\right)$$

where S is the reported sales expressed in € mn

⇒ This adjustment has the effect to reduce the default correlation and then the risk-weighted assets

The IRB formulas

Risk-weighted assets for corporate, sovereign, and bank exposures

Foundation IRB (FIRB)

- EAD is the amount of the claim
- For off-balance sheet items, the bank uses the CCF values of the SA approach.
- PD is estimated by the bank
- LGD is set to 45% for senior claims and 75% for subordinated claims
- M is set to 2.5 years

Advanced IRB (AIRB)

- For off-balance sheet items, the bank may estimate its own internal measures of CCF
- PD is estimated by the bank
- LGD may be estimated by the bank
- M is the weighted average time of the cash flows, with a one-year floor and a five-year cap

The IRB formulas

Risk-weighted assets for corporate, sovereign, and bank exposures

Example

We consider a senior debt of \$3 mn on a corporate firm. The residual maturity of the debt is equal to 2 years. We estimate the one-year probability of default at 5%

We first calculate the default correlation:

$$\rho\left(\mathrm{PD}\right) = 12\% \times \left(\frac{1 - e^{-50 \times 0.05}}{1 - e^{-50}}\right) + 24\% \times \left(1 - \frac{1 - e^{-50 \times 0.05}}{1 - e^{-50}}\right) = 12.985\%$$

We have:

$$b\left(\mathrm{PD}\right) = \left(0.11852 - 0.05478 \times \ln\left(0.05\right)\right)^2 = 0.0799$$

It follows that the maturity adjustment is equal to:

$$\varphi\left(\mathrm{M}\right) = \frac{1 + (2 - 2.5) \times 0.0799}{1 - 1.5 \times 0.0799} = 1.0908$$

The IRB formulas

Risk-weighted assets for corporate, sovereign, and bank exposures

The normalized capital charge with a one-year maturity is:

$$\begin{aligned}\mathcal{K}^{\star} &= 45\% \times \Phi\left(\frac{\Phi^{-1}(5\%) + \sqrt{12.985\%}\Phi^{-1}(99.9\%)}{\sqrt{1 - 12.985\%}}\right) - 45\% \times 5\% \\ &= 0.1055\end{aligned}$$

When the maturity is two years, we obtain:

$$\mathcal{K}^{\star} = 0.1055 \times 1.0908 = 0.1151$$

We deduce the value taken by the risk weight:

$$\mathrm{RW} = 12.5 \times 0.1151 = 143.87\%$$

It follows that the risk-weighted asset amount is equal to \$4.316 mn whereas the capital charge is \$345 287

The IRB formulas

Risk-weighted assets for corporate, sovereign, and bank exposures

Table: IRB risk weights (in %) for corporate exposures

Maturity		M = 1		M = 2.5		M = 2.5 (SME)	
LGD		45%	75%	45%	75%	45%	75%
PD (in %)	0.10	18.7	31.1	29.7	49.4	23.3	38.8
	0.50	52.2	86.9	69.6	116.0	54.9	91.5
	1.00	73.3	122.1	92.3	153.9	72.4	120.7
	2.00	95.8	159.6	114.9	191.4	88.5	147.6
	5.00	131.9	219.8	149.9	249.8	112.3	187.1
	10.00	175.8	292.9	193.1	321.8	146.5	244.2
	20.00	223.0	371.6	238.2	397.1	188.4	314.0

(*) For SME claims, sales are equal to 5 € mn

The IRB formulas

Risk-weighted assets for retail exposures

Claims can be included in the regulatory retail portfolio if they meet the following criteria:

1. The exposure must be to an individual person or to a small business
2. It satisfies the granularity criterion, meaning that no aggregate exposure to one counterpart can exceed 0.2% of the overall regulatory retail portfolio
3. The aggregated exposure to one counterparty cannot exceed 1 € mn

The IRB formulas

Risk-weighted assets for retail exposures

The maturity is set to one year:

$$\mathcal{K}^{\star} = \mathrm{LGD} \cdot \Phi \left(\frac{\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho(\mathrm{PD})}\Phi^{-1}(99.9\%)}{\sqrt{1-\rho(\mathrm{PD})}} \right) - \mathrm{LGD} \cdot \mathrm{PD}$$

- Residential mortgage exposures:

$$\rho(\mathrm{PD}) = 15\%$$

- Qualifying revolving retail exposures:

$$\rho(\mathrm{PD}) = 4\%$$

- Other retail exposures:

$$\rho(\mathrm{PD}) = 3\% \times \left(\frac{1-e^{-35\times\mathrm{PD}}}{1-e^{-35}} \right) + 16\% \times \left(1 - \frac{1-e^{-35\times\mathrm{PD}}}{1-e^{-35}} \right)$$

The IRB formulas

Risk-weighted assets for retail exposures

Table: IRB risk weights (in %) for retail exposures

		Mortgage		Revolving		Other retail	
LGD		45%	25%	45%	85%	45%	85%
PD (in %)	0.10	10.7	5.9	2.7	5.1	11.2	21.1
	0.50	35.1	19.5	10.0	19.0	32.4	61.1
	1.00	56.4	31.3	17.2	32.5	45.8	86.5
	2.00	87.9	48.9	28.9	54.6	58.0	109.5
	5.00	148.2	82.3	54.7	103.4	66.4	125.5
	10.00	204.4	113.6	83.9	158.5	75.5	142.7
	20.00	253.1	140.6	118.0	222.9	100.3	189.4

Pillar 2 – Supervisory review process

Supervisory review process (SRP)

1. Supervisory review and evaluation process (SREP)
2. Internal capital adequacy assessment process (ICAAP)

⇒ SREP defines the regulatory response to the first pillar (validation processes of internal models), whereas ICAAP addresses risks that are not captured in Pillar 1 like:

- Concentration risk and non-granular portfolios
- Default correlation
- Stressed parameters (PD and LGD)
- *Point-in-time* (PIT) versus *through the-cycle* (TTC)

Pillar 3 – Market discipline

The third pillar requires banks to publish comprehensive information about their risk management process

Since 2015, standardized templates for quantitative disclosure with a fixed format in order to facilitate the comparison between banks

The Basel III revision

For credit risk capital requirements, Basel III is close to the Basel II framework with some adjustments, which mainly concern the parameters

Remark

SA and IRB methods continue to be the two approaches for computing the capital charge for credit risk

The Basel III revision

The standardized approach

Differences between Basel II et and Basel III:

- Two methods:
  1. External credit risk assessment approach (ECRA)
  2. Standardized credit risk approach (SCRA)
- Loan-to-value ratio (LTV)

The Basel III revision
The standardized approach (ECRA)

Table: Risk weights of the SA approach (ECRA, Basel III)

Rating		AAA to AA−	A+ to A−	BBB+ to BBB−	BB+ to B−	CCC+ to C	NR
Sovereigns		0%	20%	50%	100%	150%	100%
PSE	1	20%	50%	100%	100%	150%	100%
	2	20%	50%	50%	100%	150%	50%
MDB		20%	30%	50%	100%	150%	50%
Banks	2	20%	30%	50%	100%	150%	SCRA
	2 ST	20%	20%	20%	50%	150%	SCRA
	Covered	10%	20%	20%	50%	100%	
Corporates		20%	50%	75%	100%	150%	100%
Retail*		75%					

(∗) The retail category includes revolving credits, credit cards, consumer credit loans, auto loans, student loans, etc., but not real estate exposures

The Basel III revision

The standardized approach (SCRA, banks)

The standardized credit risk approach (SCRA) must be used for all exposures to banks in two situations:

1. When the exposure is unrated
2. When external credit ratings are prohibited (e.g. in the US[12])

In this case, the bank must conduct a due diligence analysis in order to classify the exposures into three grades

- A Grade A refers to the most solid banks, whose capital exceeds the minimum regulatory capital requirements ($\mathrm{RW} = 40\%$ – 20% for short-term exposures)
- B Grade B refers to banks subject to substantial credit risk ($\mathrm{RW} = 75\%$ – 50% for short-term exposures)
- C Grade C refers to the most vulnerable banks ($\mathrm{RW} = 150\%$ – 150% for short-term exposures)

[12] The United States had abandoned in 2010 the use of commercial credit ratings after the Dodd-Frank reform

The Basel III revision

The standardized approach (SCRA, corporates)

When external credit ratings are prohibited, the risk weight of exposures to corporates is equal to 100% with two exceptions:

- A 65% risk weight is assigned to corporates, which can be considered investment grade (IG)
- For exposures to small and medium-sized enterprises, a 75% risk weight can be applied if the exposure can be classified in the retail category and 85% for the others

The Basel III revision

The standardized approach (ECRA, real estate)

Table: Risk weights of the SA approach (ECRA, Basel III)

Residential real estate			Commercial real estate		
Cash flows	ND	D	Cash flows	ND	D
$\mathrm{LTV} \leq 50$	20%	30%	$\mathrm{LTV} \leq 60$	$\min(60\%, \mathrm{RW}_C)$	70%
$50 < \mathrm{LTV} \leq 60$	25%	35%			
$60 < \mathrm{LTV} \leq 80$	30%	45%	$60 < \mathrm{LTV} \leq 80$	RW_C	90%
$80 < \mathrm{LTV} \leq 90$	40%	60%	$\mathrm{LTV} > 80$	RW_C	110%
$90 < \mathrm{LTV} \leq 100$	50%	75%			
$\mathrm{LTV} > 100$	70%	105%			

The Basel III revision

The standardized approach (ECRA, real estate)

Definition

The loan-to-value (LTV) ratio is the ratio of a loan to the value of an asset purchased

Example

If one borrows \$100 000 to purchase a house of \$150 000, the LTV ratio is $100\,000/150\,000$ or 66.67%

This ratio is extensively used in English-speaking countries (e.g. the United States) to measure the risk of the loan

In continental Europe, the risk of home property loans is measured by the ability of the borrower to repay the capital and service his debt, meaning that the risk of the loan is generally related to the income of the borrower

The Basel III revision

The standardized approach

For off-balance sheet items, credit conversion factors (CCF) have been revised. They can take the values 10%, 20%, 40%, 50% and 100%. This is a more granular scale without the possibility to set the CCF to 0%

The Basel III revision

The internal ratings-based approach

The methodology of the IRB approach does not change with respect to Basel II, since the formulas are the same except the correlation parameter for bank exposures:

$$\rho(\mathrm{PD}) = 15\% \times \left(\frac{1 - e^{-50 \times \mathrm{PD}}}{1 - e^{-50}} \right) + 30\% \times \left(\frac{1 - \left(1 - e^{-50 \times \mathrm{PD}}\right)}{1 - e^{-50}} \right)$$

Other changes

- For banks and large corporates, only the FIRB approach can be used
- In the AIRB approach, the estimated parameters of PD and LGD are subject to some input floors[a]
- The default values of the LGD parameter are 75% for subordinated claims, 45% for senior claims on financial institutions and 40% for senior claims on corporates in the FIRB approach

[a] For example, the minimum PD is set to 5 bps for corporate and bank exposures

Exposure at default

Definition

The exposure at default "*for an on-balance sheet or off-balance sheet item is defined as the expected gross exposure of the facility upon default of the obligor*"

$\Rightarrow \text{EAD}$ corresponds to the gross notional in the case of a loan or a credit

The big issue concerns off-balance sheet items, such as revolving lines of credit, credit cards or home equity lines of credit (HELOC)

Exposure at default

At the default time $\boldsymbol{\tau}$, we have:

$$\mathrm{EAD}\left(\boldsymbol{\tau} \mid t\right) = B\left(t\right) + \mathrm{CCF} \cdot \left(L\left(t\right) - B\left(t\right)\right)$$

where:

- $B\left(t\right)$ is the outstanding balance (or current drawn) at time t
- $L\left(t\right)$ is the current undrawn limit of the credit facility
- CCF is the credit conversion factor
- $L\left(t\right) - B\left(t\right)$ is the current undrawn or the amount that the debtor is able to draw upon in addition to the current drawn $B\left(t\right)$

We deduce that:

$$\mathrm{CCF} = \frac{\mathrm{EAD}\left(\boldsymbol{\tau} \mid t\right) - B\left(t\right)}{L\left(t\right) - B\left(t\right)}$$

Exposure at default

Let us consider the off-balance sheet item i that has defaulted. We have:

$$\mathrm{CCF}_i\left(\boldsymbol{\tau}_i - t\right) = \frac{B_i\left(\boldsymbol{\tau}_i\right) - B_i\left(t\right)}{L_i\left(t\right) - B_i\left(t\right)}$$

At time $\boldsymbol{\tau}_i$, we observe the default of Asset i and the corresponding exposure at default, which is equal to the outstanding balance $B_i\left(\boldsymbol{\tau}_i\right)$

$\Rightarrow$ We have to choose a date $t < \boldsymbol{\tau}_i$ to observe $B_i\left(t\right)$ and $L_i\left(t\right)$ in order to calculate the CCF

Estimation of CCF is difficult because it is sensitive to the date t

Loss given default

Loss given default versus recovery rate

- The recovery is the percentage of the notional on the defaulted debt that can be recovered
- In the Basel framework, the recovery rate is not explicitly used, and the concept of loss given default is preferred for measuring the credit portfolio loss
- We have:

$$\mathrm{LGD} \geq 1 - \boldsymbol{\mathcal{R}}$$

Loss given default

Example

We consider a bank that is lending \$100 mn to a corporate firm. We assume that the firm defaults at one time and, the bank recovers \$60 mn and the litigation costs are equal to \$5 mn

We deduce that the recovery rate is equal to:

$$\mathcal{R} = \frac{60}{100} = 60\%$$

In order to recover \$60 mn, the bank has incurred some operational and litigation costs. In this case, the bank has lost \$40 mn plus \$5 mn, implying that the loss given default is equal to:

$$\mathrm{LGD} = \frac{40 + 5}{100} = 45\%$$

Loss given default

Relationship between $\mathcal{R}$ and LGD

We have:

$$\mathrm{LGD} = 1 - \mathcal{R} + c$$

where c is the litigation cost (expressed in %)

Loss given default

Two approaches for modeling LGD:

1. The first approach considers that LGD is a random variable, whose probability distribution has to be estimated:

$$\mathrm{LGD} \sim \mathbf{F}(x)$$

2. The second approach consists in estimating the conditional expectation:

$$\mathbb{E}[\mathrm{LGD}] = \mathbb{E}[\mathrm{LGD} \mid X_1 = x_1, \ldots, X_m = x_m] = g(x_1, \ldots, x_m)$$

where $(X_1, \ldots, X_m)$ are the risk factors that impact LGD

Remark

We recall that the loss given default in the Basel IRB formulas does not correspond to the random variable, but to its expectation $\mathbb{E}[\mathrm{LGD}]$. Therefore, only the mean $\mathbb{E}[\mathrm{LGD}]$ is important for Pillar 1

$\Rightarrow$ Pillar 2 uses the entire probability distribution $\mathbf{F}(x)$ and the condition expectation under stressed conditions

Loss given default

Stochastic modeling (parametric distribution)

Beta distribution

The beta distribution $\mathcal{B}(\alpha, \beta)$ has the following pdf:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathfrak{B}(\alpha, \beta)}$$

where $\mathfrak{B}(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$. The mean and the variance are:

$$\mu(X) = \mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$$

and:

$$\sigma^2(X) = \operatorname{var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

When α and β are greater than 1, the distribution has one mode $x_{\mathrm{mode}} = (\alpha-1)/(\alpha+\beta-2)$

Loss given default

Stochastic modeling (parametric distribution)

Several shapes:

- $\mathcal{B}(1,1) \sim \mathcal{U}_{[0,1]}$, $\mathcal{B}(\infty,\infty) \sim \delta_{0.5}([0,1])$, $\mathcal{B}(\alpha,0) \sim \mathcal{B}(1)$ and $\mathcal{B}(0,\beta) \sim \mathcal{B}(0)$
- If $\alpha = \beta$, the distribution is symmetric around $x = 0.5$; we have a bell curve when the two parameters α and β are higher than 1, and a **U**-shape curve when the two parameters α and β are lower than 1
- If $\alpha > \beta$, the skewness is negative and the distribution is left-skewed, if $\alpha < \beta$, the skewness is positive and the distribution is right-skewed

Loss given default

Stochastic modeling (parametric distribution)

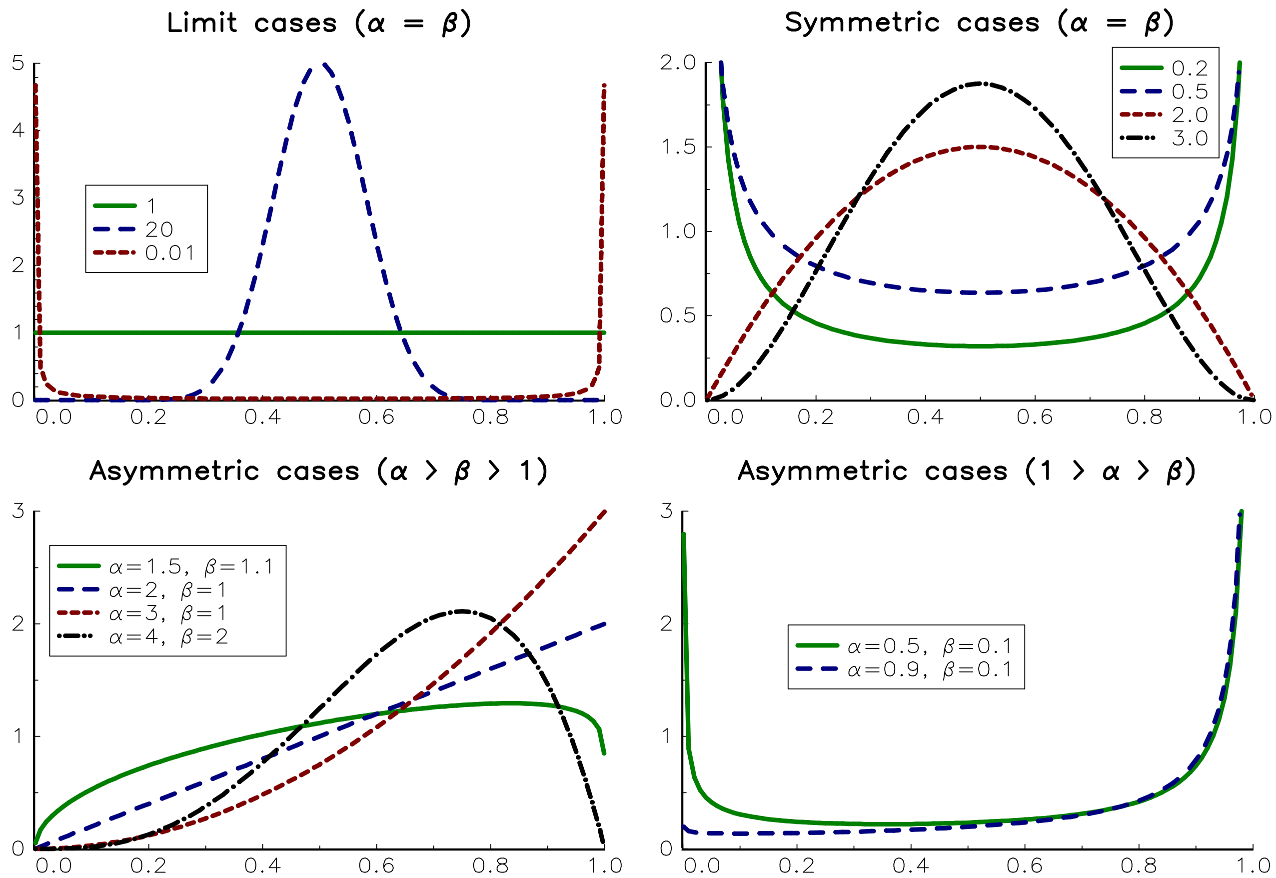


Figure: Probability density function of the beta distribution $\mathcal{B}(\alpha, \beta)$

Loss given default

Stochastic modeling (parametric distribution)

Method of moments (HFRM, Section 10.1.3, page 628)

We have:

$$\hat{\alpha}_{\mathrm{MM}} = \frac{\hat{\mu}_{\mathrm{LGD}}^2 \left(1 - \hat{\mu}_{\mathrm{LGD}}\right)}{\hat{\sigma}_{\mathrm{LGD}}^2} - \hat{\mu}_{\mathrm{LGD}}$$

and:

$$\hat{\beta}_{\mathrm{MM}} = \frac{\hat{\mu}_{\mathrm{LGD}} \left(1 - \hat{\mu}_{\mathrm{LGD}}\right)^2}{\hat{\sigma}_{\mathrm{LGD}}^2} - \left(1 - \hat{\mu}_{\mathrm{LGD}}\right)$$

Maximum likelihood estimation (HFRM, Section 10.1.2, page 614)

$$\begin{aligned} \left(\hat{\alpha}_{\mathrm{ML}}, \hat{\beta}_{\mathrm{ML}}\right) &= \arg\max \boldsymbol{\ell}\left(\alpha, \beta\right) \\ &= \arg\max \left(\alpha - 1\right) \sum_{i=1}^{n} \ln y_i + \left(b - 1\right) \sum_{i=1}^{n} \ln\left(1 - y_i\right) - n \ln \mathfrak{B}\left(\alpha, \beta\right) \end{aligned}$$

Loss given default

Stochastic modeling (parametric distribution)

Example

We consider the following sample of losses given default: 68%, 90%, 22%, 45%, 17%, 25%, 89%, 65%, 75%, 56%, 87%, 92% and 46%

We obtain $\hat{\mu}_{\mathrm{LGD}} = 59.77\%$ and $\hat{\sigma}_{\mathrm{LGD}} = 27.02\%$. Using the method of moments, the estimated parameters are $\hat{\alpha}_{\mathrm{MM}} = 1.37$ and $\hat{\beta}_{\mathrm{MM}} = 0.92$

Using a **numerical optimization** method, we have $\hat{\alpha}_{\mathrm{ML}} = 1.84$ and $\hat{\beta}_{\mathrm{ML}} = 1.25$. See HFRM on page 619 for the statistical inference:

Table: Results of the maximum likelihood estimation

Parameter	Estimate	Standard error	t-statistic	p-value
α	1.8356	0.6990	2.6258	0.0236
β	1.2478	0.4483	2.7834	0.0178

Loss given default

Stochastic modeling (parametric distribution)

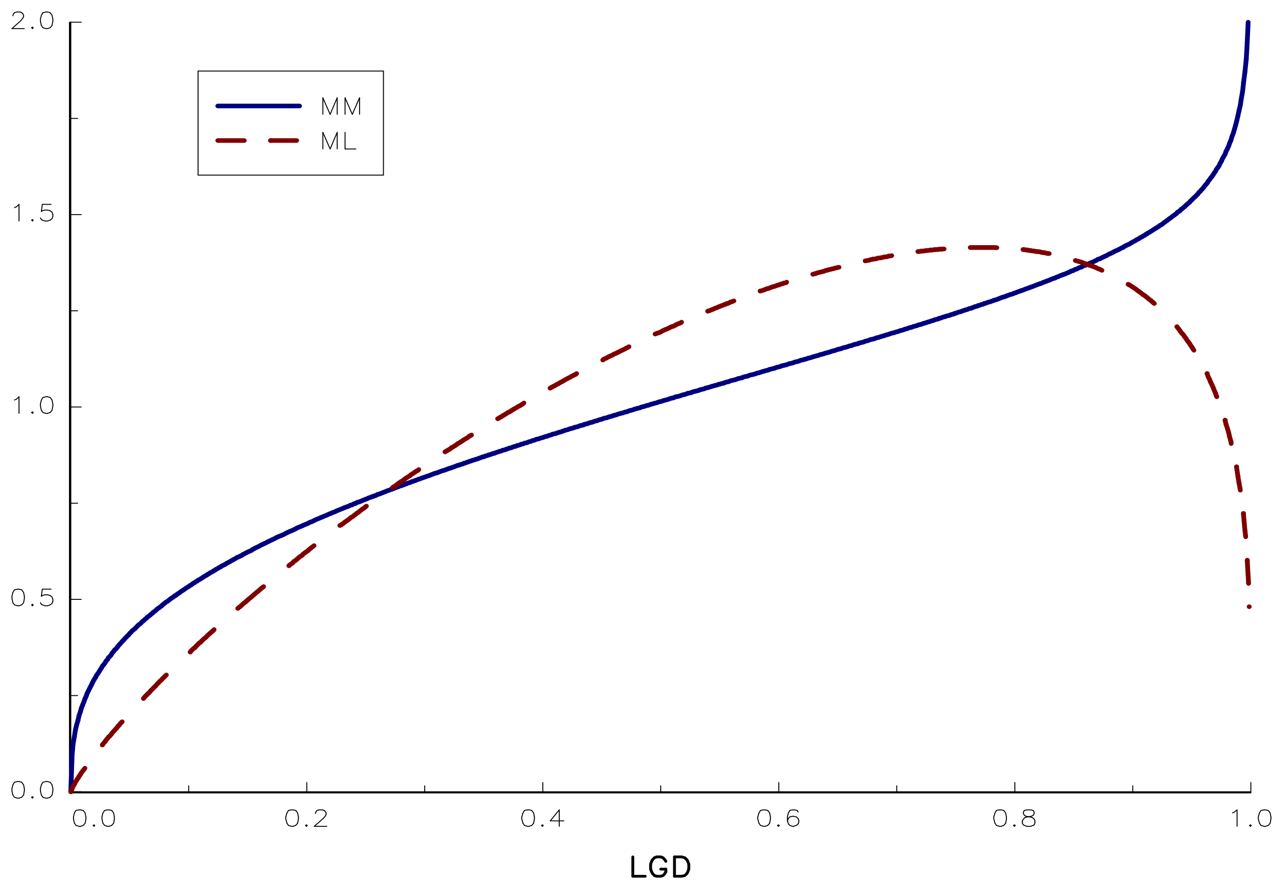


Figure: Calibration of the beta distribution

Loss given default

Stochastic modeling (non-parametric distribution)

The limit case of the beta distribution's **U**-shaped is the Bernoulli distribution:

LGD	0%	100%
Probability	$(1-\mu_{\text{LGD}})$	μ_{LGD}

$\Rightarrow$ Extension to the empirical distribution or histogram

Example

We consider the following empirical distribution of LGD:

LGD (in %)	0	10	20	25	30	40	50	60	70	75	80	90	100
$\hat{p}$ (in %)	1	2	10	25	10	2	0	2	10	25	10	2	1

Loss given default

Stochastic modeling (non-parametric distribution)

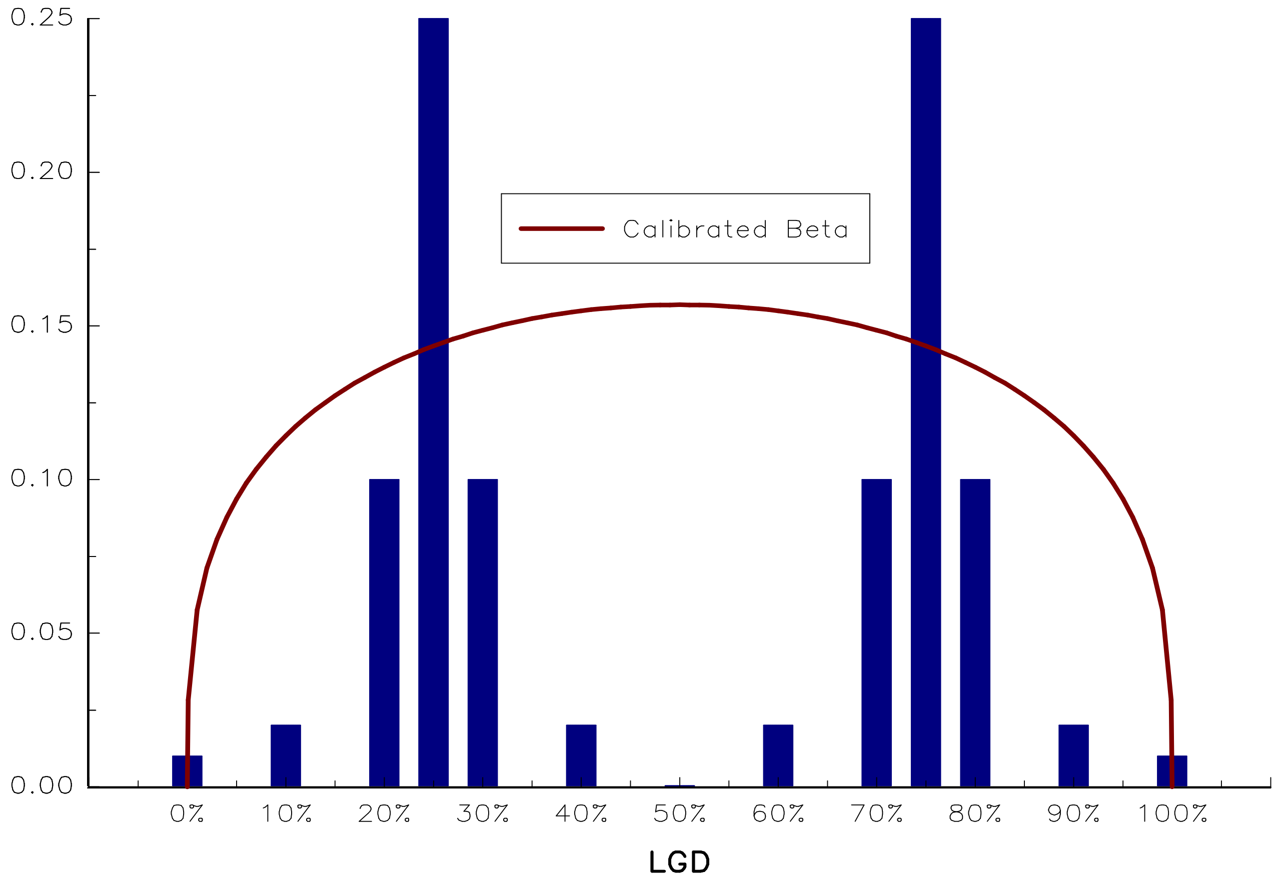


Figure: Calibration of a bimodal LGD distribution

Loss given default

The case of non-granular portfolios

Example

We consider a credit portfolio of 10 loans, whose loss is equal to:

$$L = \sum_{i=1}^{10} \mathrm{EaD}_i \cdot \mathrm{LGD}_i \cdot \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

where T_i is equal to 5 years, EaD_i is equal to \$1 000 and the default time $\boldsymbol{\tau}_i$ is exponential with the following intensity parameter λ_i:

i	1	2	3	4	5	6	7	8	9	10
λ_i (in bps)	10	10	25	25	50	100	250	500	500	1 000

The loss given default LGD_i is given by the previous empirical distribution:

LGD (in %)	0	10	20	25	30	40	50	60	70	75	80	90	100
$\hat{p}$ (in %)	1	2	10	25	10	2	0	2	10	25	10	2	1

Loss given default
The case of non-granular portfolios

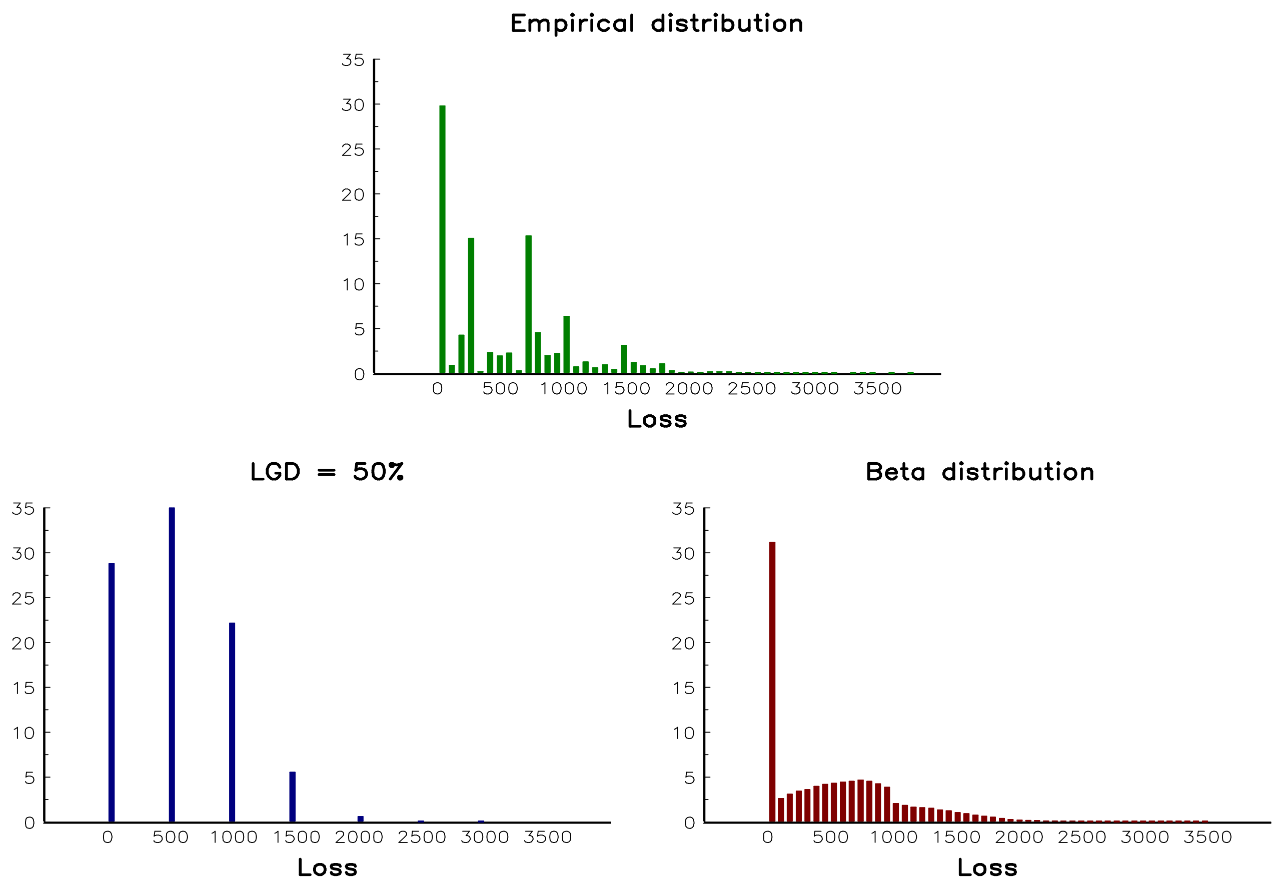


Figure: Loss frequency in % of the three LGD models

Loss given default

The case of non-granular portfolios

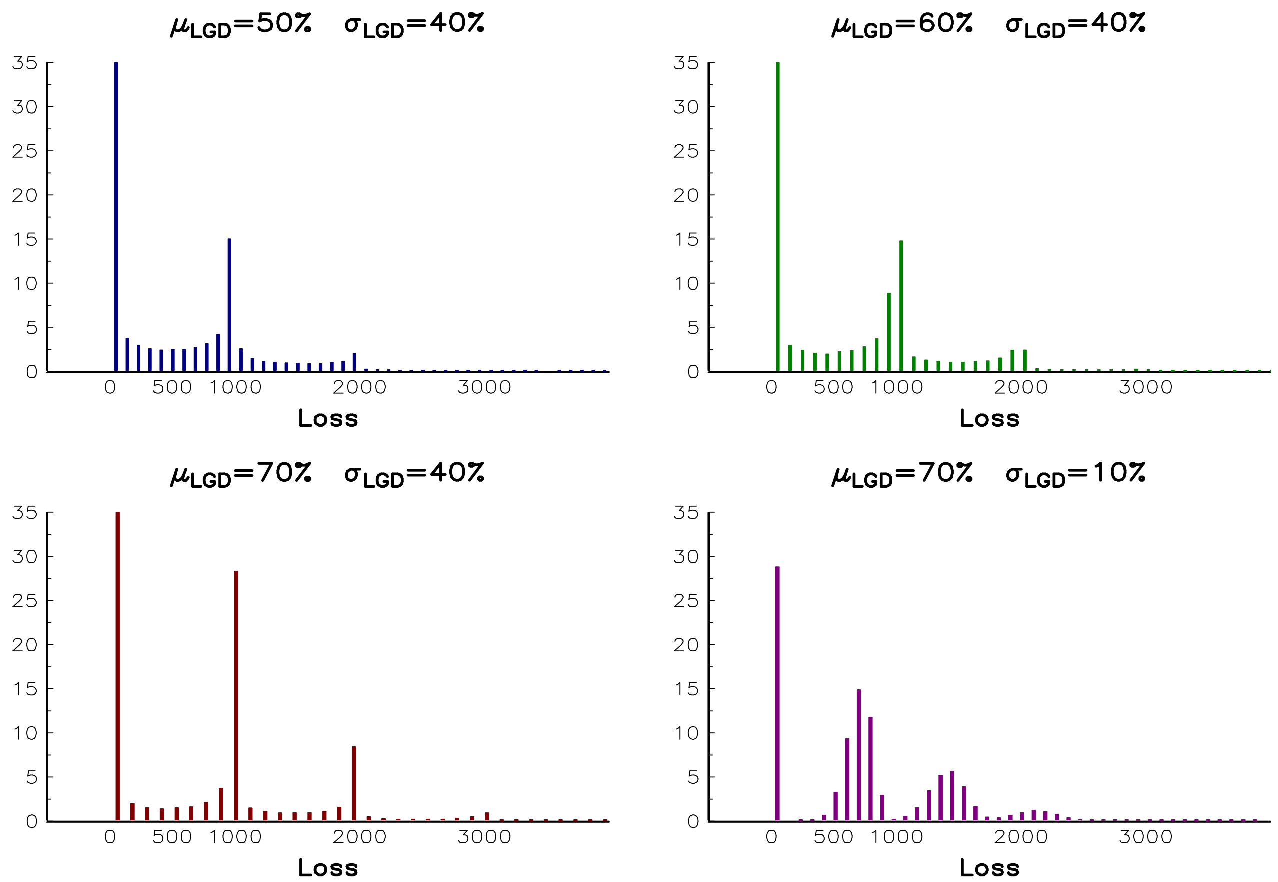


Figure: Loss frequency in % for different values of μ_{LGD} and σ_{LGD}

Loss given default

The case of granular portfolios

Expression of the portfolio loss

We recall that:

$$L = \sum_{i=1}^{n} \mathrm{EAD}_i \cdot \mathrm{LGD}_i \cdot \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

If the portfolio is fined grained, we have:

$$\mathbb{E}\left[L \mid X\right] = \sum_{i=1}^{n} \mathrm{EAD}_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot p_i\left(X\right)$$

We deduce that the distribution of the portfolio loss does not depend on the random variables LGD_i, but on their expected values $\mathbb{E}\left[\mathrm{LGD}_i\right]$. Therefore, we can replace the previous expression of the portfolio loss by:

$$L = \sum_{i=1}^{n} \mathrm{EAD}_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

Loss given default
Economic modeling

The third version of Moody's LossCalc considers seven factors that are grouped in three major categories:

1. factors external to the issuer: geography, industry, credit cycle stage
2. factors specific to the issuer: distance-to-default, probability of default (or leverage for private firms)
3. factors specific to the debt issuance: debt type, relative standing in capital structure, collateral

Once the factors are identified, we must estimate the LGD model:

$$\mathrm{LGD} = f\left(X_1, \ldots, X_m\right)$$

where $X_1, \ldots, X_m$ are the m factors, and f is a non-linear function

We apply a logit transformation and estimate the model using linear regression or quantile regression (see HFRM, Section 14.2.3, page 909) $\Rightarrow$ This approach will be studied in Lecture 11 dedicated to stress testing and scenario analysis

Probability of default

Three approaches:

- Survival function
- Transition probability matrix
- Structural models

Survival function

Let $\boldsymbol{\tau}$ be a default (or survival) time. The survival function is defined as follows:

$$\mathbf{S}(t) = \Pr\{\boldsymbol{\tau} > t\} = 1 - \mathbf{F}(t)$$

where $\mathbf{F}$ is the cumulative distribution function. We deduce that:

$$f(t) = -\frac{\partial\,\mathbf{S}(t)}{\partial\,t}$$

We define the hazard function $\lambda(t)$ as the instantaneous default rate given that the default has not occurred before t:

$$\lambda(t) = \lim_{\mathrm{d}t \to 0^+} \frac{\Pr\{t \leq \boldsymbol{\tau} \leq t + \mathrm{d}t \mid \boldsymbol{\tau} \geq t\}}{\mathrm{d}t}$$

We deduce that:

$$\begin{aligned}
\lambda(t) &= \lim_{\mathrm{d}t \to 0^+} \frac{\Pr\{t \leq \boldsymbol{\tau} \leq t + \mathrm{d}t\}}{\mathrm{d}t} \cdot \frac{1}{\Pr\{\boldsymbol{\tau} \geq t\}} \\
&= \frac{f(t)}{\mathbf{S}(t)} = -\frac{\partial_t\,\mathbf{S}(t)}{\mathbf{S}(t)} = -\frac{\partial\,\ln \mathbf{S}(t)}{\partial\,t}
\end{aligned}$$

Survival function

The survival function can then be rewritten with respect to the hazard function and we have:

$$\mathbf{S}(t) = e^{-\int_0^t \lambda(s)\,\mathrm{d}s}$$

Table: Common survival functions

Model	$\mathbf{S}(t)$	$\lambda(t)$
Exponential	$\exp(-\lambda t)$	λ
Weibull	$\exp(-\lambda t^{\gamma})$	$\lambda \gamma t^{\gamma-1}$
Log-normal	$1 - \Phi(\gamma \ln(\lambda t))$	$\gamma t^{-1} \phi(\gamma \ln(\lambda t)) / (1 - \Phi(\gamma \ln(\lambda t)))$
Log-logistic	$1 / \left(1 + \lambda t^{\frac{1}{\gamma}}\right)$	$\lambda \gamma^{-1} t^{\frac{1}{\gamma}} / \left(t + \lambda t^{1+\frac{1}{\gamma}}\right)$
Gompertz	$\exp(\lambda(1 - e^{\gamma t}))$	$\lambda \gamma \exp(\gamma t)$
Cox	$\mathbf{S}(t) = e^{-\exp\left(\beta^{\top} x\right) \int_0^t \lambda_0(s)\,\mathrm{d}s}$	$\lambda_0(t) \exp\left(\beta^{\top} x\right)$

Exponential survival time

We note $\boldsymbol{\tau} \sim \mathcal{E}(\lambda)$ and we have:

$$\mathbf{S}(t) = e^{-\lambda t}$$

Main properties

1. The mean residual life $\mathbb{E}[\boldsymbol{\tau} \mid \boldsymbol{\tau} \geq t]$ is constant
2. It satisfies the **lack of memory property** (LMP):

$$\Pr\{\boldsymbol{\tau} \geq t + u \mid \boldsymbol{\tau} \geq t\} = \Pr\{\boldsymbol{\tau} \geq u\}$$

   or equivalently $\mathbf{S}(t+u) = \mathbf{S}(t)\,\mathbf{S}(u)$
3. The probability distribution of $n \cdot \boldsymbol{\tau}_{1:n}$ is the same as probability distribution of $\boldsymbol{\tau}_i$

Piecewise exponential model

We have:

$$\lambda(t)=\sum_{m=1}^{M}\lambda_m\cdot\mathbb{1}\left\{t_{m-1}^{\star}<t\leq t_m^{\star}\right\}=\lambda_m \qquad \text{if } t\in\left]t_{m-1}^{\star},t_m^{\star}\right]$$

where $t_m^{\star}$ are the knots of the function $\left(t_0^{\star}=0,\ t_{M+1}^{\star}=\infty\right)$. For $t\in\left]t_{m-1}^{\star},t_m^{\star}\right]$, the expression of the survival function becomes:

$$\mathbf{S}(t)=\exp\left(-\sum_{k=1}^{m-1}\lambda_k\left(t_k^{\star}-t_{k-1}^{\star}\right)-\lambda_m\left(t-t_{m-1}^{\star}\right)\right)=\mathbf{S}\left(t_{m-1}^{\star}\right)e^{-\lambda_m\left(t-t_{m-1}^{\star}\right)}$$

It follows that the density function is equal to:

$$f(t)=\lambda_m\exp\left(-\sum_{k=1}^{m-1}\lambda_k\left(t_k^{\star}-t_{k-1}^{\star}\right)-\lambda_m\left(t-t_{m-1}^{\star}\right)\right)$$

We verify that:

$$\frac{f(t)}{\mathbf{S}(t)}=\lambda_m \qquad \text{if } t\in\left]t_{m-1}^{\star},t_m^{\star}\right]$$

Piecewise exponential model

Example

We consider three set of parameters $\{(t_m^\star, \lambda_m), m = 1, \ldots, M\}$:

$$\begin{aligned}
\{(1, 1\%), (2, 1.5\%), (3, 2\%), (4, 2.5\%), (\infty, 3\%)\} & \quad \text{for } \lambda_1(t) \\
\{(1, 10\%), (2, 7\%), (5, 5\%), (7, 4.5\%), (\infty, 6\%)\} & \quad \text{for } \lambda_2(t) \\
\lambda_3(t) = 4\% & \quad \text{for } \lambda_3(t)
\end{aligned}$$

Piecewise exponential model

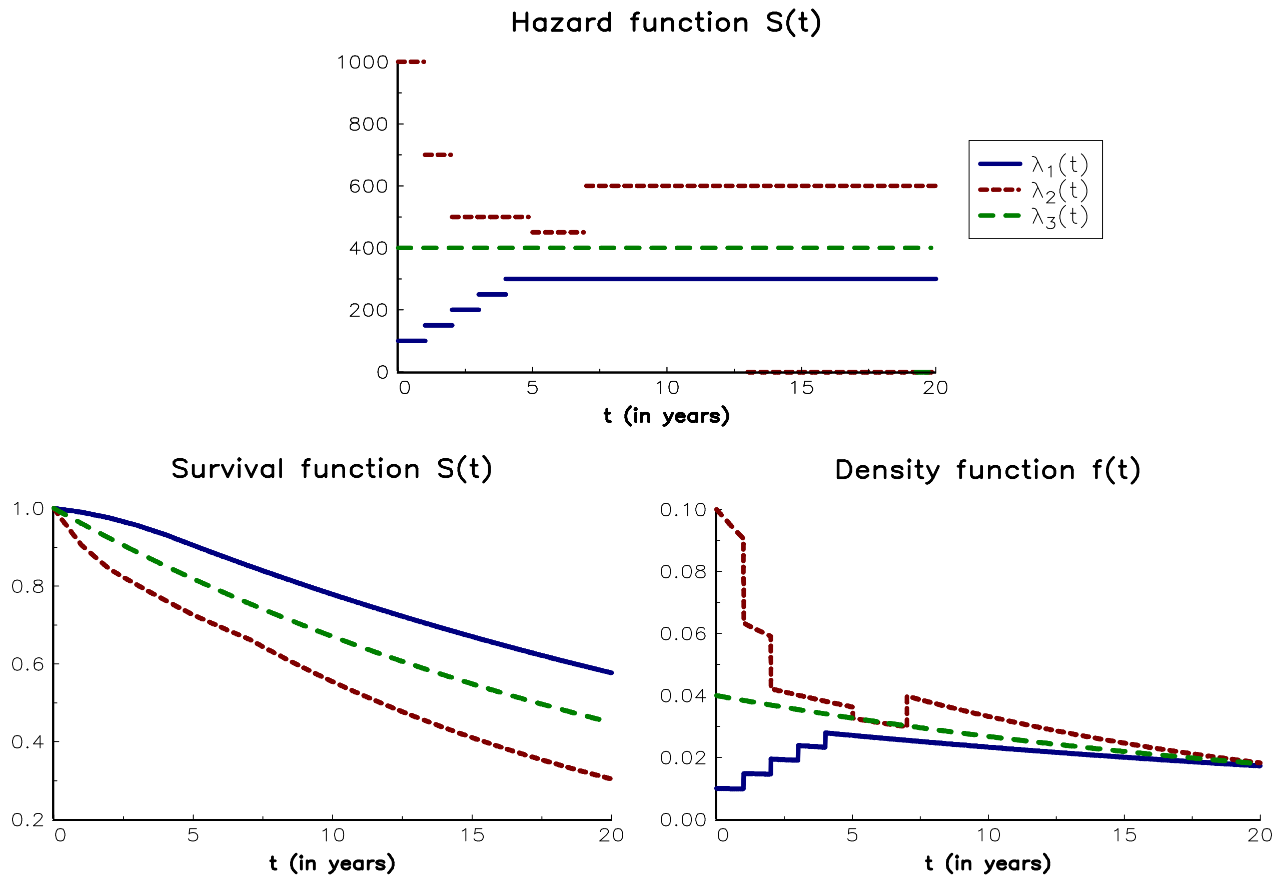


Figure: Example of the piecewise exponential model

Piecewise exponential model

Estimation methods:

- Non-linear least squares regression
- Kaplan-Meier estimation (non-parametric approach)
- Bootstrap

Bootstrap method

1. We first estimate the parameter λ_1 for the earliest maturity Δt_1
2. Assuming that $\left(\hat{\lambda}_1, \ldots, \hat{\lambda}_{i-1}\right)$ have been estimated, we calculate $\hat{\lambda}_i$ for the next maturity Δt_i
3. We iterate step 2 until the last maturity Δt_m

$\Rightarrow$ This algorithm is used for calibrating the credit curve of CDS spreads

Piecewise exponential model

Example

We consider three credit curves, whose CDS spreads expressed in bps are given in the table below. We assume that the recovery rate $\mathcal{R}$ is set to 40%

Table: Calibrated piecewise exponential model from CDS prices

	Credit curve			Bootstrap solution		
Maturity (in years)	#1	#2	#3	#1	#2	#3
1	50	50	350	83.3	83.3	582.9
3	60	60	370	110.1	110.1	637.5
5	70	90	390	140.3	235.0	702.0
7	80	115	385	182.1	289.6	589.4
10	90	125	370	194.1	241.9	498.5

Transition probability matrix

Definition

We consider a time-homogeneous Markov chain $\mathfrak{R}$, whose transition matrix is $P = (p_{i,j})$. We note $\mathcal{S} = \{1, 2, \ldots, K\}$ the state space of the chain and $p_{i,j}$ is the probability that the entity migrates from rating i to rating j. The matrix P satisfies the following properties:

- $\forall i, j \in \mathcal{S}$, $p_{i,j} \geq 0$;
- $\forall i \in \mathcal{S}$, $\sum_{j=1}^{K} p_{i,j} = 1$.

In credit risk, we generally assume that K is the absorbing state (or the default state), implying that any entity which has reached this state remains in this state ($p_{K,K} = 1$)

Transition probability matrix

Table: Example of credit migration matrix (in %)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	92.82	6.50	0.56	0.06	0.06	0.00	0.00	0.00
AA	0.63	91.87	6.64	0.65	0.06	0.11	0.04	0.00
A	0.08	2.26	91.66	5.11	0.61	0.23	0.01	0.04
BBB	0.05	0.27	5.84	87.74	4.74	0.98	0.16	0.22
BB	0.04	0.11	0.64	7.85	81.14	8.27	0.89	1.06
B	0.00	0.11	0.30	0.42	6.75	83.07	3.86	5.49
CCC	0.19	0.00	0.38	0.75	2.44	12.03	60.71	23.50
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Transition probability matrix

Let $\mathfrak{R}(t)$ be the value of the state at time t. We define $p(s,i;t,j)$ as the probability that the entity reaches the state j at time t given that it has reached the state i at time s:

$$p(s,i;t,j) = \Pr\{\mathfrak{R}(t) = j \mid \mathfrak{R}(s) = i\} = p_{i,j}^{(t-s)}$$

This is the Markov property

The n-step transition probability is defined as:

$$p_{i,j}^{(n)} = \Pr\{\mathfrak{R}(t+n) = j \mid \mathfrak{R}(t) = i\}$$

and we note $P^{(n)} = \left(p_{i,j}^{(n)}\right)$ the associated n-step transition matrix

Transition probability matrix

For $n = 2$, we obtain:

$$\begin{aligned}
p_{i,j}^{(2)} &= \Pr \{\Re(t+2) = j \mid \Re(t) = i\} \\
&= \sum_{k=1}^{K} \Pr \{\Re(t+2) = j, \Re(t+1) = k \mid \Re(t) = i\} \\
&= \sum_{k=1}^{K} \Pr \{\Re(t+2) = j \mid \Re(t+1) = k\} \cdot \Pr \{\Re(t+1) = k \mid \Re(t) = i\} \\
&= \sum_{k=1}^{K} p_{i,k} \cdot p_{k,j}
\end{aligned}$$

Transition probability matrix

Chapman-Kolmogorov (forward) equation

We have (scalar form):

$$p_{i,j}^{(n+m)} = \sum_{k=1}^{K} p_{i,k}^{(n)} \cdot p_{k,j}^{(m)} \qquad \forall n, m > 0$$

or (matrix form):

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)}$$

with the convention $P^{(0)} = I_K$

We deduce that:

$$P^{(n)} = P^n$$

and:

$$p\left(t, i; t+n, j\right) = p_{i,j}^{(n)} = \mathbf{e}_i^\top P^n \mathbf{e}_j$$

Transition probability matrix

$$
\begin{aligned}
p^{(2)}_{\text{AAA},\text{AAA}} &= p_{\text{AAA},\text{AAA}} \times p_{\text{AAA},\text{AAA}} + p_{\text{AAA},\text{AA}} \times p_{\text{AA},\text{AAA}} + p_{\text{AAA},\text{A}} \times p_{\text{A},\text{AAA}} + \\
&\quad p_{\text{AAA},\text{BBB}} \times p_{\text{BBB},\text{AAA}} + p_{\text{AAA},\text{BB}} \times p_{\text{BB},\text{AAA}} + p_{\text{AAA},\text{B}} \times p_{\text{B},\text{AAA}} + \\
&\quad p_{\text{AAA},\text{CCC}} \times p_{\text{CCC},\text{AAA}} \\
&= 0.9283^2 + 0.0650 \times 0.0063 + 0.0056 \times 0.0008 + \\
&\quad 0.0006 \times 0.0005 + 0.0006 \times 0.0004 \\
&= 86.1970\%
\end{aligned}
$$

Transition probability matrix

Table: Two-year transition probability matrix P^2 (in %)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	86.20	12.02	1.47	0.18	0.11	0.01	0.00	0.00
AA	1.17	84.59	12.23	1.51	0.18	0.22	0.07	0.02
A	0.16	4.17	84.47	9.23	1.31	0.51	0.04	0.11
BBB	0.10	0.63	10.53	77.66	8.11	2.10	0.32	0.56
BB	0.08	0.24	1.60	13.33	66.79	13.77	1.59	2.60
B	0.01	0.21	0.61	1.29	11.20	70.03	5.61	11.03
CCC	0.29	0.04	0.68	1.37	4.31	17.51	37.34	38.45
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Transition probability matrix

Table: Five-year transition probability matrix P^5 (in %)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	69.23	23.85	5.49	0.96	0.31	0.12	0.02	0.03
AA	2.35	66.96	24.14	4.76	0.86	0.62	0.13	0.19
A	0.43	8.26	68.17	17.34	3.53	1.55	0.18	0.55
BBB	0.24	1.96	19.69	56.62	13.19	5.32	0.75	2.22
BB	0.17	0.73	5.17	21.23	40.72	20.53	2.71	8.74
B	0.07	0.47	1.73	4.67	16.53	44.95	5.91	25.68
CCC	0.38	0.24	1.37	2.92	7.13	18.51	9.92	59.53
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Transition probability matrix

We note $\pi_i^{(n)}$ the probability of the state i at time n:

$$\pi_i^{(n)} = \Pr \{\mathfrak{R}(n) = i\}$$

and $\pi^{(n)} = \left(\pi_1^{(n)}, \ldots, \pi_K^{(n)}\right)$ the probability distribution. By construction, we have:

$$\pi^{(n+1)} = P^\top \pi^{(n)}$$

The Markov chain $\mathfrak{R}$ admits a stationary distribution $\pi^\star$ if $\pi^\star = P^\top \pi^\star$:

$$\lim_{n\to\infty} p_{k,i}^{(n)} = \pi_i^\star$$

We can interpret $\pi_i^\star$ as the average duration spent by the Markov chain $\mathfrak{R}$ in the state i

Transition probability matrix

Average return period of a Markov chain

Let $\mathcal{T}_i$ be the return period of state i:

$$\mathcal{T}_i = \inf \left\{ n : \mathfrak{R}(n) = i \mid \mathfrak{R}(0) = i \right\}$$

The average return period is then equal to:

$$\mathbb{E}\left[\mathcal{T}_i\right] = \frac{1}{\pi_i^\star}$$

Transition probability matrix

Survival function

Survival function

Since K is the default state, the survival function $\mathbf{S}_i(t)$ of a firm whose initial rating is the state i is given by:

$$\begin{aligned}
\mathbf{S}_i(t) &= 1 - \Pr\{\mathfrak{R}(t) = K \mid \mathfrak{R}(0) = i\} \\
&= 1 - \mathbf{e}_i^\top P^t \mathbf{e}_K
\end{aligned}$$

Transition probability matrix

Survival function

Estimation of the piecewise exponential model

In the piecewise exponential model, the survival function is

$$\mathbf{S}(t) = \mathbf{S}\left(t_{m-1}^{\star}\right) e^{-\lambda_m \left(t - t_{m-1}^{\star}\right)}$$

for $t \in \left] t_{m-1}^{\star}, t_m^{\star} \right]$. We deduce that $\mathbf{S}\left(t_m^{\star}\right) = \mathbf{S}\left(t_{m-1}^{\star}\right) e^{-\lambda_m \left(t_m^{\star} - t_{m-1}^{\star}\right)}$, implying that:

$$\ln \mathbf{S}\left(t_m^{\star}\right) = \ln \mathbf{S}\left(t_{m-1}^{\star}\right) - \lambda_m \left(t_m^{\star} - t_{m-1}^{\star}\right)$$

and:

$$\lambda_m = \frac{\ln \mathbf{S}\left(t_{m-1}^{\star}\right) - \ln \mathbf{S}\left(t_m^{\star}\right)}{t_m^{\star} - t_{m-1}^{\star}}$$

Transition probability matrix

Survival function

Estimation of the piecewise exponential model

It is then straightforward to estimate the piecewise hazard function from a transition probability matrix:

- The knots of the piecewise function are the years $m \in \mathbb{N}^*$
- For each initial rating i, the hazard function $\lambda_i(t)$ is defined as:

$$\lambda_i(t) = \lambda_{i,m} \qquad \text{if } t \in \left]m-1, m\right]$$

where:

$$\begin{aligned} \lambda_{i,m} &= \frac{\ln \mathbf{S}_i(m-1) - \ln \mathbf{S}_i(m)}{m-(m-1)} \\ &= \ln \left(\frac{1 - \mathbf{e}_i^\top P^{m-1} \mathbf{e}_K}{1 - \mathbf{e}_i^\top P^m \mathbf{e}_K} \right) \end{aligned}$$

and $P^0 = I$

Transition probability matrix

Survival function

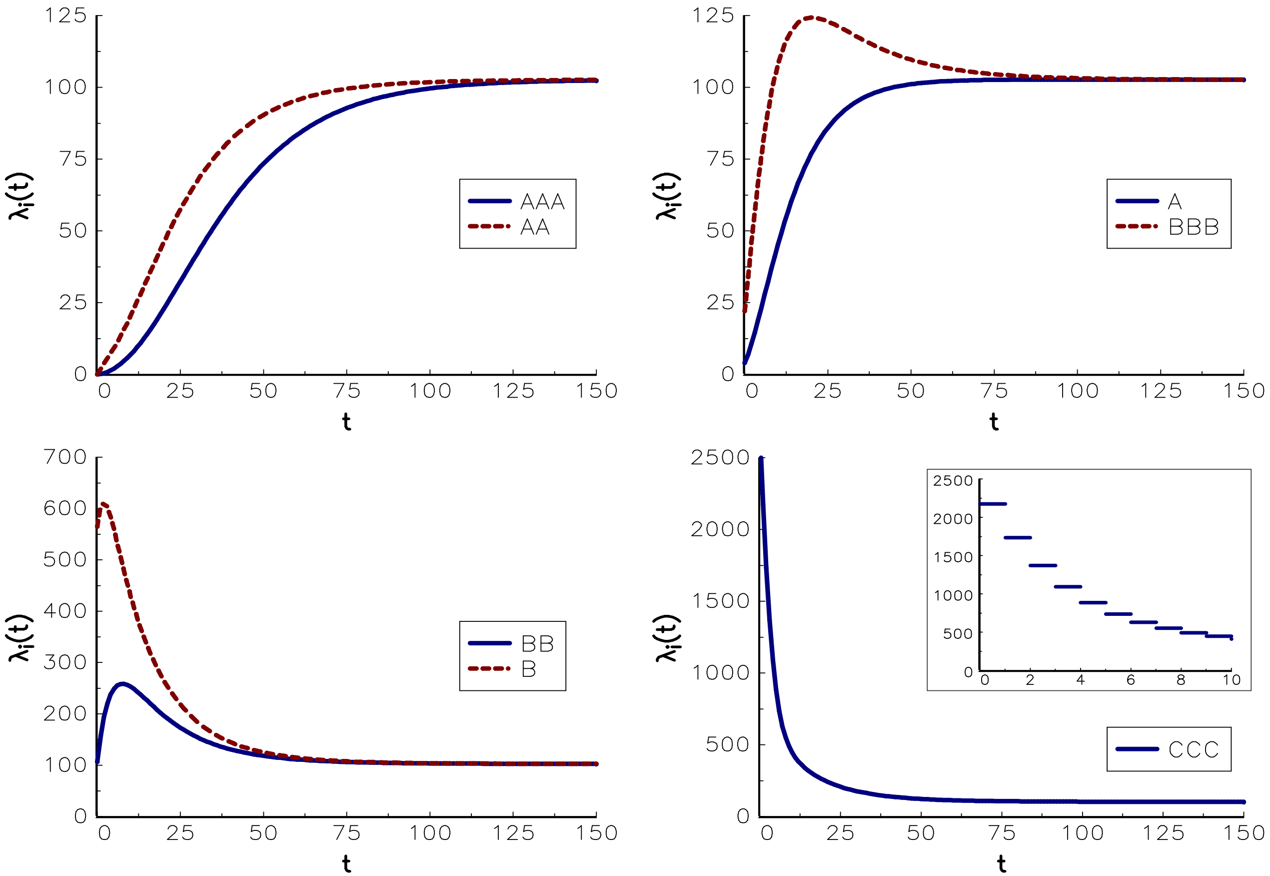


Figure: Estimated hazard function $\lambda_i(t)$ from the credit migration matrix

Transition probability matrix
Survival function

Why the hazard function of all the ratings converges to the same level, which is equal to 102.63 bps?

In the long run, the initial rating has no impact on the survival function:

Conditional probability distribution $\Rightarrow$ Unconditional probability distribution

We deduce that the annual default rate is exactly equal to 1.0263%

Transition probability matrix

Continuous-time modeling

Definition

The transition matrix $P(s;t)$ is defined as follows:

$$P_{i,j}(s;t) = p(s,i;t,j) = \Pr\left\{\mathfrak{R}(t) = j \mid \mathfrak{R}(s) = i\right\}$$

where $s \in \mathbb{R}_+$ and $t \in \mathbb{R}_+$. Assuming that the Markov chain is time-homogenous, we have $P(t) = P(0;t)$

Markov generator

The Markov generator is defined by the matrix $\Lambda = (\lambda_{i,j})$ where $\lambda_{i,j} \geq 0$ for all $i \neq j$ and $\lambda_{i,i} = -\sum_{j\neq i}^{K} \lambda_{i,j}$. In this case, the transition matrix satisfies the following relationship:

$$P(t) = \exp(t\Lambda)$$

where $\exp(A)$ is the matrix exponential of A.

Transition probability matrix

Continuous-time modeling

Probabilistic interpretation of Λ

If we assume that the probability of jumping from rating i to rating j in a short time period Δt is proportional to Δt, we have:

$$p(t, i; t + \Delta t, j) = \lambda_{i,j} \Delta t$$

The matrix form of this equation is $P(t; t + \Delta t) = \Lambda \, \Delta t$. We deduce that:

$$P(t + \Delta t) = P(t) \, P(t; t + \Delta t) = P(t) \, \Lambda \, \Delta t$$

and:

$$\mathrm{d}P(t) = P(t) \, \Lambda \, \mathrm{d}t$$

Because we have $\exp(\mathbf{0}) = I$, we obtain the solution $P(t) = \exp(t\Lambda)$

$\lambda_{i,j}$ can be interpreted as the instantaneous transition rate of jumping from rating i to rating j

Transition probability matrix

Matrix exponential (HFRM, Appendix A.1.1.3, page 1034)

Let $f(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. The matrix exponential of the matrix A is equal to:

$$B = e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

whereas the matrix logarithm of A is the matrix B such that $e^B = A$ and we note $B = \ln A$

Let A and B be two $n \times n$ square matrices. Using the Taylor expansion, we can show that $f\left(A^\top\right) = f(A)^\top$, $Af(A) = f(A)A$ and $f\left(B^{-1}AB\right) = B^{-1}f(A)B$. It follows that $e^{A^\top} = \left(e^A\right)^\top$ and $e^{B^{-1}AB} = B^{-1}e^AB$. If $AB = BA$, we can also prove that $Ae^B = e^BA$ and $e^{A+B} = e^Ae^B = e^Be^A$

Remark

Algorithms for computing matrix functions (e^A, $\ln A$, A^x, $\sqrt{A}$, $\cos A$, etc.) are available in programming languages (matlab, gauss, python, etc.)

Transition probability matrix

Continuous-time modeling

Example

We consider a rating system with three states: A (good rating), B (bad rating) and D (default). The Markov generator is equal to:

$$\Lambda = \begin{pmatrix} -0.30 & 0.20 & 0.10 \\ 0.15 & -0.40 & 0.25 \\ 0.00 & 0.00 & 0.00 \end{pmatrix}$$

The one-year transition probability matrix is equal to:

$$P(1) = e^{\Lambda} = \begin{pmatrix} 75.16\% & 14.17\% & 10.67\% \\ 10.63\% & 68.07\% & 21.30\% \\ 0.00\% & 0.00\% & 100.00\% \end{pmatrix}$$

Transition probability matrix

Continuous-time modeling

For the two-year maturity, we get:

$$P(2) = e^{2\Lambda} = \begin{pmatrix} 58.00\% & 20.30\% & 21.71\% \\ 15.22\% & 47.85\% & 36.93\% \\ 0.00\% & 0.00\% & 100.00\% \end{pmatrix}$$

We verify that $P(2) = P(1)^2$. This derives from the property of the matrix exponential:

$$P(t) = e^{t\Lambda} = \left(e^{\Lambda}\right)^t = P(1)^t$$

Transition probability matrix

Continuous-time modeling

The one-month transition probability matrix is equal to:

$$P\left(\frac{1}{12}\right) = e^{\frac{1}{12}\Lambda} = \begin{pmatrix} 97.54\% & 1.62\% & 0.84\% \\ 1.21\% & 96.73\% & 2.05\% \\ 0.00\% & 0.00\% & 100.00\% \end{pmatrix}$$

Remark

Another way to compute the one-month transition probability matrix is to use the matrix exponent function:

$$P\left(\frac{1}{12}\right) = P(1)^{\frac{1}{12}}$$

Transition probability matrix

Continuous-time modeling

Let $\hat{P}(t)$ be the empirical transition matrix for a given t. We can estimate the Markov generator:

$$\hat{\Lambda} = \frac{1}{t} \ln \left(\hat{P}(t) \right)$$

Table: Markov generator $\hat{\Lambda}$ (in bps)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	−747.49	703.67	35.21	3.04	6.56	**-0.79**	**-0.22**	0.02
AA	67.94	−859.31	722.46	51.60	2.57	10.95	4.92	**-1.13**
A	7.69	245.59	−898.16	567.70	53.96	20.65	**-0.22**	2.80
BBB	5.07	21.53	650.21	−1352.28	557.64	85.56	16.08	16.19
BB	4.22	10.22	41.74	930.55	−2159.67	999.62	97.35	75.96
B	**-0.84**	11.83	30.11	8.71	818.31	−1936.82	539.18	529.52
CCC	25.11	**-2.89**	44.11	84.87	272.05	1678.69	−5043.00	2941.06
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The matrix $\hat{\Lambda}$ does not verify the Markov conditions $\hat{\lambda}_{i,j} \geq 0$ for all $i \neq j$

Transition probability matrix

Continuous-time modeling

Israel *et al.* (2001) propose two estimators to obtain a valid generator:

1. The first approach consists in adding the negative values back into the diagonal values:
$$\left\{\begin{array}{ll} \bar{\lambda}_{i,j} = \max\left(\hat{\lambda}_{i,j},0\right) & i \neq j \\ \bar{\lambda}_{i,i} = \hat{\lambda}_{i,i} + \sum_{j\neq i} \min\left(\hat{\lambda}_{i,j},0\right) & \end{array}\right.$$

2. In the second method, we carry forward the negative values on the matrix entries which have the correct sign:
$$\left\{\begin{array}{l} G_i = \left|\hat{\lambda}_{i,i}\right| + \sum_{j\neq i} \max\left(\hat{\lambda}_{i,j},0\right) \\ B_i = \sum_{j\neq i} \max\left(-\hat{\lambda}_{i,j},0\right) \\ \tilde{\lambda}_{i,j} = \left\{\begin{array}{ll} 0 & \text{if } i \neq j \text{ and } \hat{\lambda}_{i,j} < 0 \\ \hat{\lambda}_{i,j} - B_i \left|\hat{\lambda}_{i,j}\right| / G_i & \text{if } G_i > 0 \\ \hat{\lambda}_{i,j} & \text{if } G_i = 0 \end{array}\right. \end{array}\right.$$

Transition probability matrix

Continuous-time modeling

Table: Markov generator $\tilde{\Lambda}$ (in bps)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	−747.99	703.19	35.19	3.04	6.55	0.00	0.00	0.02
AA	67.90	−859.88	721.98	51.57	2.57	10.94	4.92	0.00
A	7.69	245.56	−898.27	567.63	53.95	20.65	0.00	2.80
BBB	5.07	21.53	650.21	−1352.28	557.64	85.56	16.08	16.19
BB	4.22	10.22	41.74	930.55	−2159.67	999.62	97.35	75.96
B	0.00	11.83	30.10	8.71	818.14	−1937.24	539.06	529.40
CCC	25.10	0.00	44.10	84.84	271.97	1678.21	−5044.45	2940.22
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table: 207-day transition probability matrix (in %)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	95.85	3.81	0.27	0.03	0.04	0.00	0.00	0.00
AA	0.37	95.28	3.90	0.34	0.03	0.06	0.02	0.00
A	0.04	1.33	95.12	3.03	0.33	0.12	0.00	0.02
BBB	0.03	0.14	3.47	92.75	2.88	0.53	0.09	0.11
BB	0.02	0.06	0.31	4.79	88.67	5.09	0.53	0.53
B	0.00	0.06	0.17	0.16	4.16	89.84	2.52	3.08
CCC	0.12	0.01	0.23	0.45	1.45	7.86	75.24	14.64
D	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Transition probability matrix

Continuous-time modeling

Remark

The continuous-time framework is more flexible when modeling credit risk. For instance, the expression of the survival function becomes:

$$\mathbf{S}_i(t) = \Pr\{\mathfrak{R}(t) = K \mid \mathfrak{R}(0) = i\} = 1 - \mathbf{e}_i^\top \exp(t\Lambda)\,\mathbf{e}_K$$

We can therefore calculate the probability density function in an easier way:

$$f_i(t) = -\partial_t\,\mathbf{S}_i(t) = \mathbf{e}_i^\top \Lambda \exp(t\Lambda)\,\mathbf{e}_K$$

Transition probability matrix

Continuous-time modeling

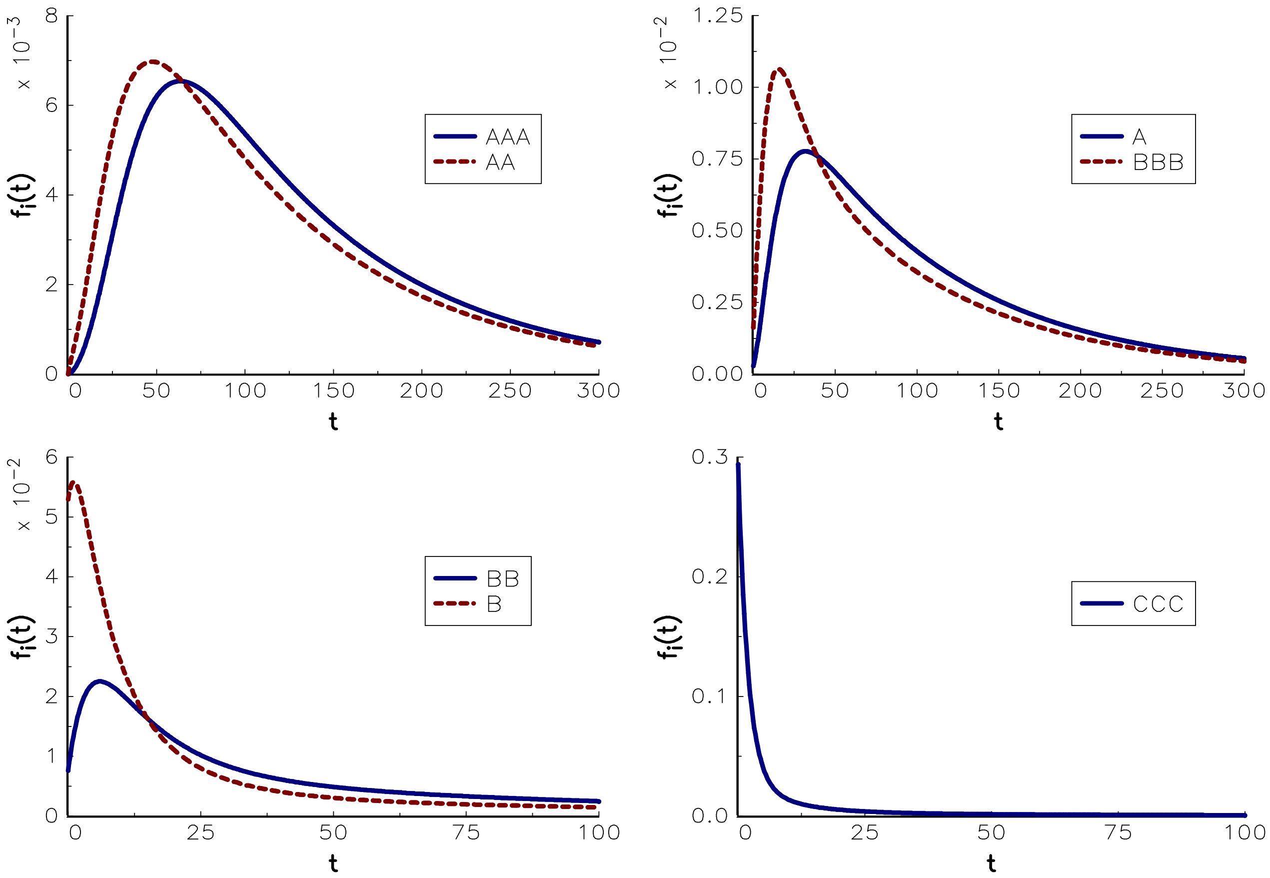


Figure: Probability density function $f_i(t)$ of S&P ratings

Structural models

Two main models:

- Merton (1974)
- Black and Cox (1976)

Two main implementations:

- KMV
- CreditGrades

Other topics

Pillar 1

- Exposure at default
- Expected loss given default
- Probability of default

Pillar 2

- Random loss given default
- Default correlation
- Granularity

Internal model

- Exposure at default
- Random loss given default
- Probability of default
- Default correlation
- Granularity

Default correlation

Two approaches:

- Copula models
- Factor models

$\Rightarrow$ Same concept

Default correlation
The copula model

Let $\mathbf{S}$ be the survival function of the random vector $(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_n)$, we can show that $\mathbf{S}$ admits a copula representation:

$$\mathbf{S}(t_1, \ldots, t_n) = \mathbf{C}(\mathbf{S}_1(t_1), \ldots, \mathbf{S}_n(t_n))$$

where $\mathbf{S}_i$ is the survival function of $\boldsymbol{\tau}_i$ and $\mathbf{C}$ is the survival copula associated to $\mathbf{S}$

Default correlation

The copula function of the Basel model

In the Basel mode, the (normalized) asset value of the i^{th} firm is $Z_i \sim \mathcal{N}(0,1)$ and the default occurs when Z_i is below a non-stochastic barrier B_i:

$$D_i = 1 \Leftrightarrow Z_i \leq B_i = \Phi^{-1}(p_i)$$

We recall that $Z_i = \sqrt{\rho} X + \sqrt{1-\rho}\varepsilon_i$ where $X \sim \mathcal{N}(0,1)$ is the systematic risk factor and $\varepsilon_i \sim \mathcal{N}(0,1)$ is the specific risk factor, and the conditional default probability is equal to:

$$p_i(X) = \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}X}{\sqrt{1-\rho}}\right)$$

If we introduce the time dimension, we obtain:

$$p_i(t) = \Pr\{\boldsymbol{\tau}_i \leq t\} = 1 - S_i(t)$$

and:

$$p_i(t, X) = \Phi\left(\frac{\Phi^{-1}(1 - \mathbf{S}_i(t)) - \sqrt{\rho}X}{\sqrt{1-\rho}}\right)$$

where $S_i(t)$ is the survival function of the i^{th} firm

Default correlation

The copula function of the Basel model

$Z = (Z_1, \ldots, Z_n) \sim \mathcal{N}(\mathbf{0}_n, \mathbb{C}_n(\rho))$ with:

$$\mathbb{C}_n(\rho) = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}$$

It follows that the **joint default probability** is:

$$\begin{aligned} p_{1,\ldots,n} &= \Pr\{D_1 = 1, \ldots, D_n = 1\} = \Pr\{Z_1 \leq B_1, \ldots, Z_n \leq B_n\} \\ &= \Phi(B_1, \ldots, B_n; \mathbb{C}_n(\rho)) \end{aligned}$$

Since we have $B_i = \Phi^{-1}(p_i)$, we deduce that:

$$p_{1,\ldots,n} = \Phi\left(\Phi^{-1}(p_1), \ldots, \Phi^{-1}(p_n); \mathbb{C}_n(\rho)\right)$$

The Basel copula between default probabilities is the Normal copula with a constant correlation matrix

Default correlation

The copula function of the Basel model

If we consider the dependence between the survival times, we have:

$$\begin{aligned}
\mathbf{S}(t_1, \ldots, t_n) &= \Pr\{\boldsymbol{\tau}_1 > t_1, \ldots, \boldsymbol{\tau}_n > t_n\} \\
&= \Pr\left\{Z_1 > \Phi^{-1}(p_1(t_1)), \ldots, Z_n > \Phi^{-1}(p_n(t_n))\right\} \\
&= \Pr\{\Phi(Z_1) > p_1(t_1), \ldots, \Phi(Z_n) > p_n(t_n)\} \\
&= \Pr\{\Phi(Z_1) \leq 1 - p_1(t_1), \ldots, \Phi(Z_n) \leq 1 - p_n(t_n)\} \\
&= \mathbf{C}(1 - p_1(t_1), \ldots, 1 - p_n(t_n); \mathbb{C}_n(\rho)) \\
&= \mathbf{C}(\mathbf{S}_1(t_1), \ldots, \mathbf{S}_n(t_n); \mathbb{C}_n(\rho))
\end{aligned}$$

The Basel copula between default times is the Normal copula with a constant correlation matrix

Default correlation

Extension to other copula functions

From an industrial point of view, only two copula functions are used and tractable:

1. The Normal copula
2. The Student t copula

with a general correlation matrix:

$$\mathbb{C} = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,n} \\ & 1 & & \vdots \\ & & \ddots & \rho_{n-1,n} \\ & & & 1 \end{pmatrix}$$

⇒ In practice, we use a structural correlation matrix (HFRM, pages 221-225)

Default correlation
The factor model

One-factor model

$$Z_i = \sqrt{\rho} X + \sqrt{1-\rho} \varepsilon_i$$

$(m+1)$-factor model

$$Z_i = \sqrt{\rho} \cdot X + \sqrt{\rho_{\mathrm{map}(i)} - \rho} \cdot X_{\mathrm{map}(i)} + \sqrt{1 - \rho_{\mathrm{map}(i)}} \cdot \varepsilon_i$$

Default correlation

Jump-to-default

How default correlations affects default times

Let $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ be two default times, whose joint survival function is $\mathbf{S}(t_1, t_2) = \mathbf{C}(\mathbf{S}_1(t_1), \mathbf{S}_2(t_2))$. We have:

$$
\begin{aligned}
\mathbf{S}_1(t \mid \boldsymbol{\tau}_2 = t^\star) &= \Pr\{\boldsymbol{\tau}_1 > t \mid \boldsymbol{\tau}_2 = t^\star\} \\
&= \partial_2 \mathbf{C}(\mathbf{S}_1(t), \mathbf{S}_2(t^\star)) \\
&= \mathbf{C}_{2|1}(\mathbf{S}_1(t), \mathbf{S}_2(t^\star)) \\
&\neq \mathbf{S}_1(t) \qquad \text{except if } \mathbf{C} = \mathbf{C}^\perp
\end{aligned}
$$

where $\mathbf{C}_{2|1}$ is the conditional copula function

$\Rightarrow$ This phenomenon is called jump-to-defaut (JTD) or spread jump

Default correlation

Jump-to-default of credit ratings

The hazard function is equal to:

$$\lambda_i(t) = \frac{f_i(t)}{\mathbf{S}_i(t)} = \frac{\mathbf{e}_i^\top \Lambda \exp(t\Lambda)\,\mathbf{e}_K}{1 - \mathbf{e}_i^\top \exp(t\Lambda)\,\mathbf{e}_K}$$

We deduce that:

$$\lambda_{i_1}(t \mid \boldsymbol{\tau}_{i_2} = t^\star) = \frac{f_{i_1}(t \mid \boldsymbol{\tau}_{i_2} = t^\star)}{\mathbf{S}_{i_1}(t \mid \boldsymbol{\tau}_{i_2} = t^\star)}$$

With the Basel copula, we have:

$$\mathbf{S}_{i_1}(t \mid \boldsymbol{\tau}_{i_2} = t^\star) = \Phi\left(\frac{\Phi^{-1}(\mathbf{S}_{i_1}(t)) - \rho\Phi^{-1}(\mathbf{S}_{i_2}(t^\star))}{\sqrt{1-\rho^2}}\right)$$

and:

$$f_{i_1}(t \mid \boldsymbol{\tau}_{i_2} = t^\star) = \phi\left(\frac{\Phi^{-1}(\mathbf{S}_{i_1}(t)) - \rho\Phi^{-1}(\mathbf{S}_{i_2}(t^\star))}{\sqrt{1-\rho^2}}\right)\frac{f_{i_1}(t)}{\sqrt{1-\rho^2}\,\phi\left(\Phi^{-1}(\mathbf{S}_{i_1}(t))\right)}$$

Default correlation

Jump-to-default of credit ratings

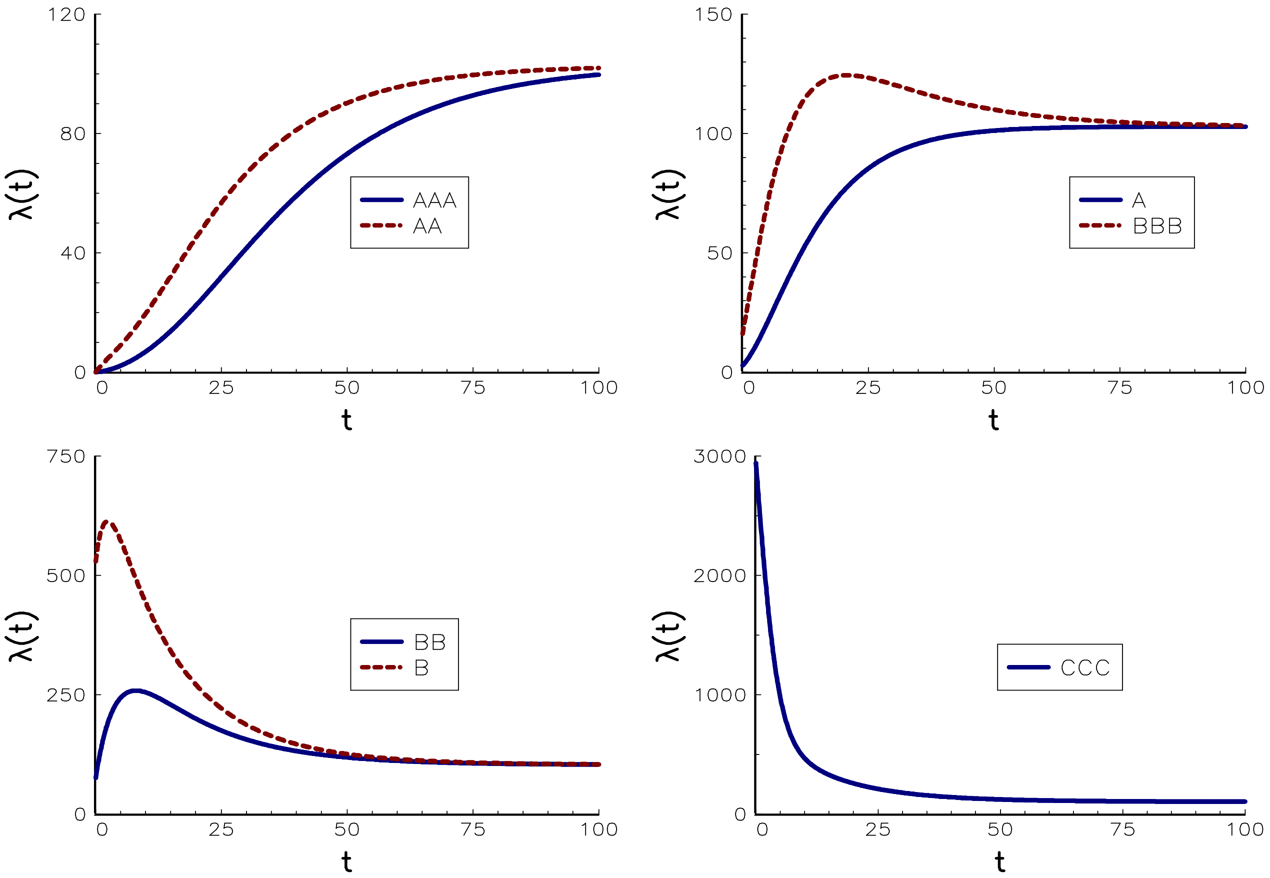


Figure: Hazard function $\lambda_i(t)$ (in bps)

Default correlation

Jump-to-default of credit ratings

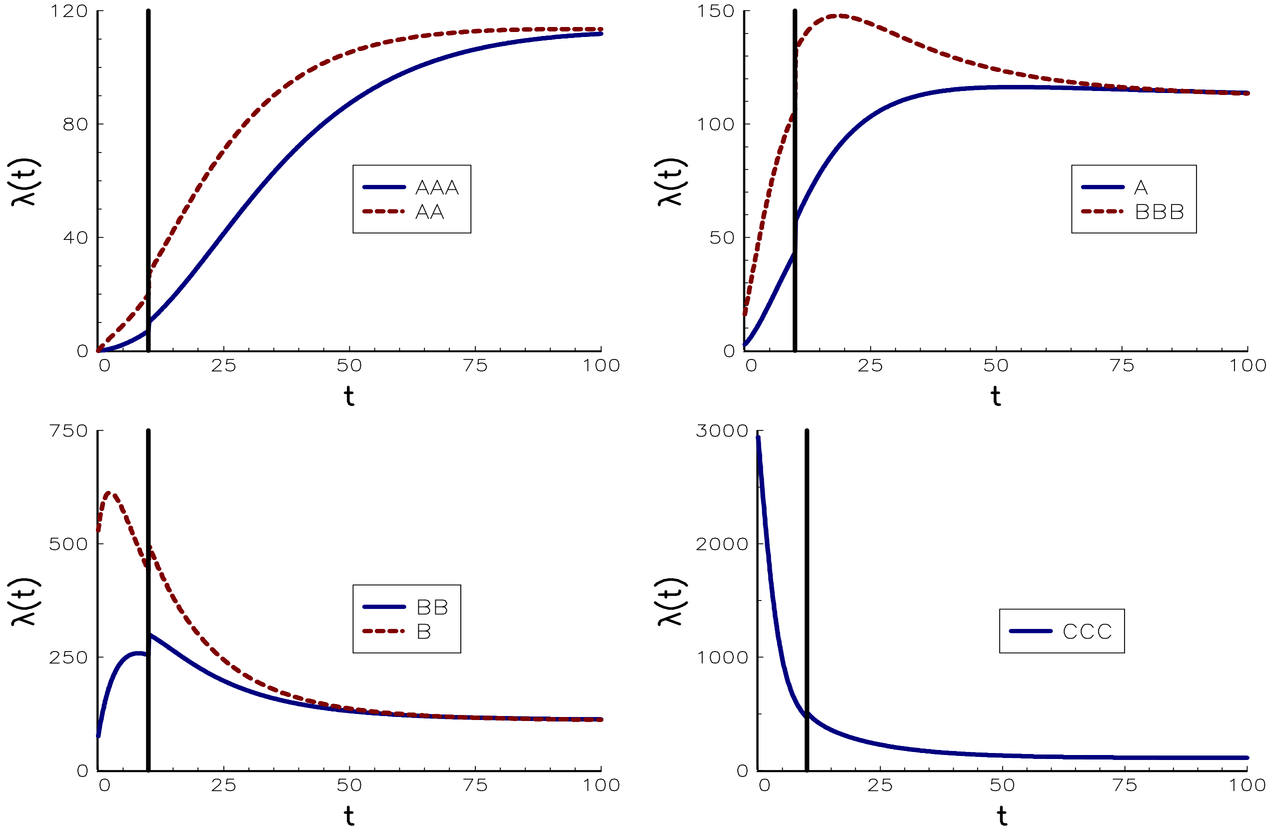


Figure: Hazard function $\lambda_i(t)$ (in bps) when a AAA-rated company defaults after 10 years ($\rho = 5\%$)

Default correlation

Jump-to-default of credit ratings

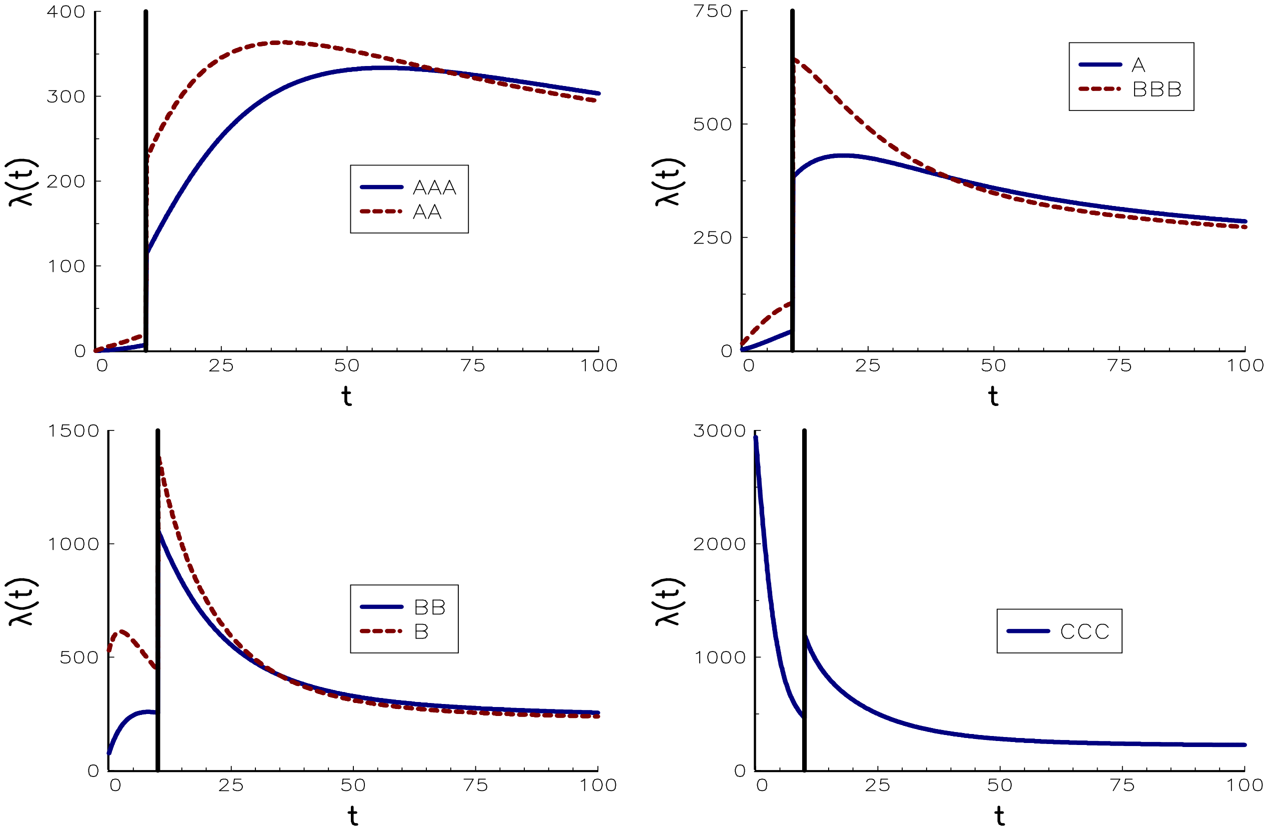


Figure: Hazard function $\lambda_i(t)$ (in bps) when a AAA-rated company defaults after 10 years ($\rho = 50\%$)

Default correlation

Jump-to-default of credit ratings

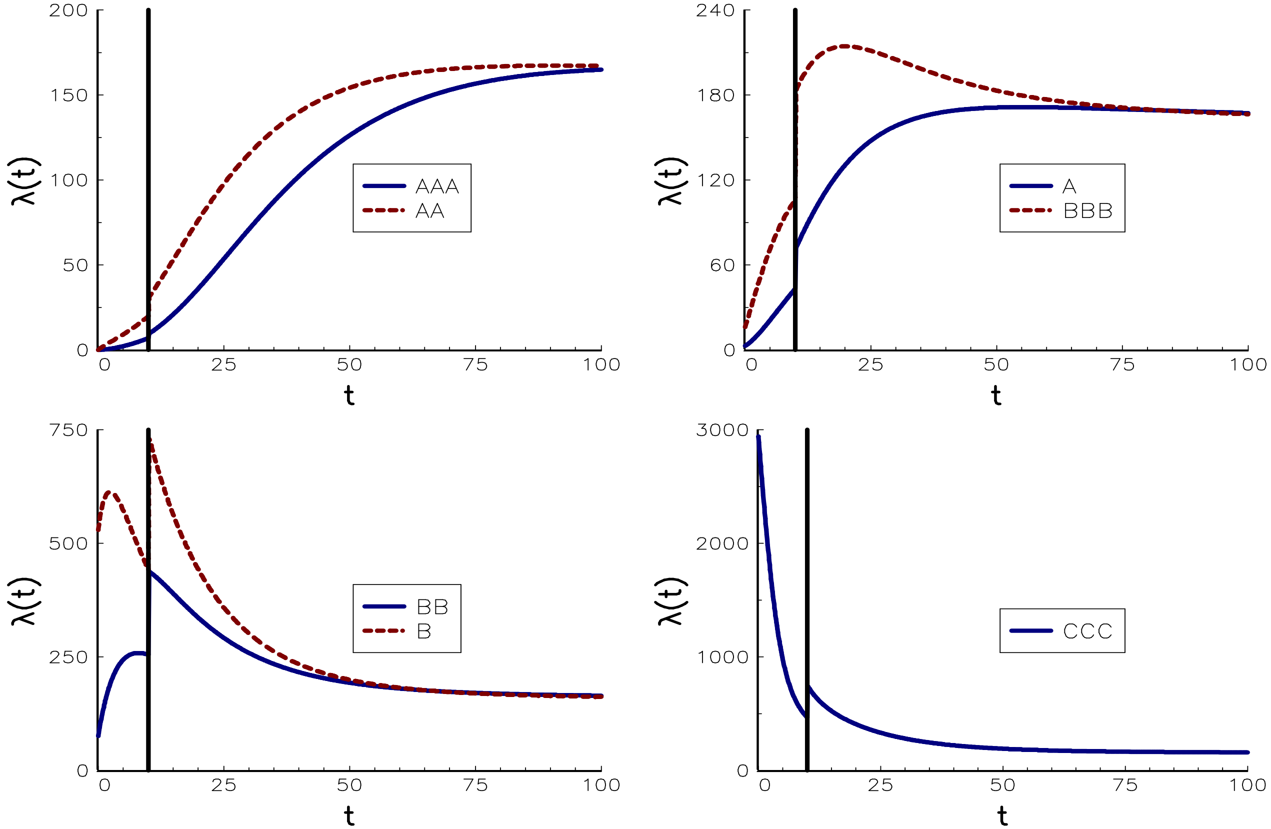


Figure: Hazard function $\lambda_i(t)$ (in bps) when a BB-rated company defaults after 10 years ($\rho = 50\%$)

Default correlation

Jump-to-default of credit ratings

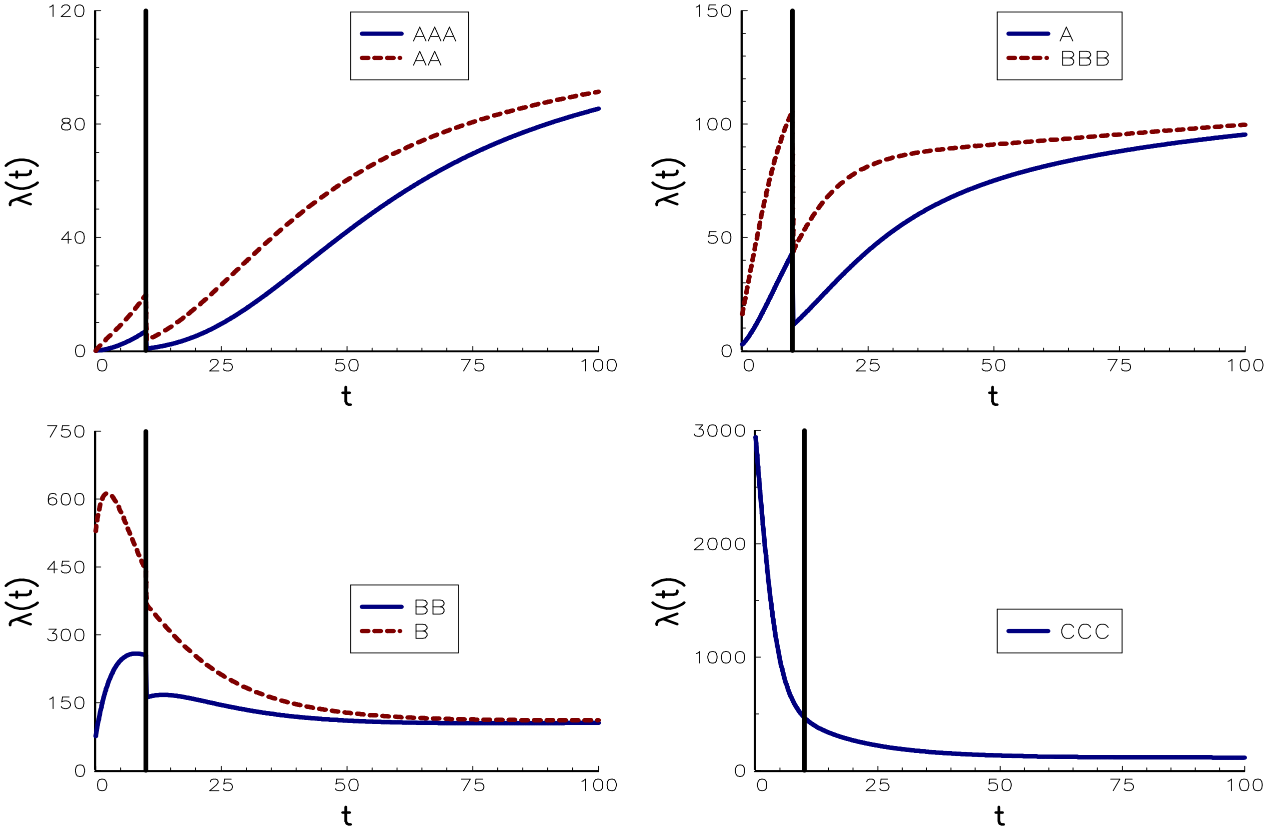


Figure: Hazard function $\lambda_i(t)$ (in bps) when a CCC-rated company defaults after 10 years ($\rho = 50\%$)

Granularity and concentration

Definition of the granularity adjustment

We recall that the portfolio loss is given by:

$$L = \sum_{i=1}^{n} \mathrm{EAD}_i \cdot \mathrm{LGD}_i \cdot \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

For an infinitely fine-grained (IFG) portfolio, we have:

$$\mathrm{VaR}_\alpha\left(w_{\mathrm{IFG}}\right) = \sum_{i=1}^{n} \mathrm{EAD}_i \cdot \mathbb{E}\left[\mathrm{LGD}_i\right] \cdot \Phi\left(\frac{\Phi^{-1}\left(\mathrm{PD}_i\right) + \sqrt{\rho}\Phi^{-1}\left(\mathrm{PD}_i\right)}{\sqrt{1-\rho}}\right)$$

However, the portfolio w cannot be fine-grained and present some concentration issues, implying that the value-at-risk is equal to the quantile α of the loss distribution:

$$\mathrm{VaR}_\alpha\left(w\right) = \mathbf{F}_L^{-1}\left(\alpha\right)$$

The granularity adjustment GA is the difference between the two risk measures:

$$\mathrm{GA} = \mathrm{VaR}_\alpha\left(w\right) - \mathrm{VaR}_\alpha\left(w_{\mathrm{IFG}}\right)$$

Granularity and concentration

The case of a perfectly concentrated portfolio

Let us consider a portfolio that is made up of one credit:

$$L = \mathrm{EAD} \cdot \mathrm{LGD} \cdot \mathbb{1}\left\{\boldsymbol{\tau} \leq T\right\}$$

It follows that:

$$\mathbf{F}_L(\ell) = \Pr\left\{\mathrm{EAD} \cdot \mathrm{LGD} \cdot \mathbb{1}\left\{\boldsymbol{\tau} \leq T\right\} \leq \ell\right\}$$

Since we have $\ell = 0 \Leftrightarrow \boldsymbol{\tau} > T$, we deduce that $\mathbf{F}_L(0) = \Pr\left\{\boldsymbol{\tau} > T\right\} = 1 - \mathrm{PD}$. If $\ell \neq 0$, we have:

$$\begin{aligned}\mathbf{F}_L(\ell) &= \mathbf{F}_L(0) + \Pr\left\{\mathrm{EAD} \cdot \mathrm{LGD} \leq \ell \mid \boldsymbol{\tau} \leq T\right\} \\ &= (1 - \mathrm{PD}) + \mathrm{PD} \cdot \mathbf{G}\left(\frac{\ell}{\mathrm{EAD}}\right)\end{aligned}$$

where $\mathbf{G}$ is the distribution function of the loss given default. The value-at-risk of this portfolio is then equal to:

$$\mathrm{VaR}_\alpha(w) = \begin{cases} \mathrm{EAD} \cdot \mathbf{G}^{-1}\left(\dfrac{\alpha + \mathrm{PD} - 1}{\mathrm{PD}}\right) & \text{if } \alpha \geq 1 - \mathrm{PD} \\ 0 & \text{otherwise} \end{cases}$$

Granularity and concentration

The case of a perfectly concentrated portfolio

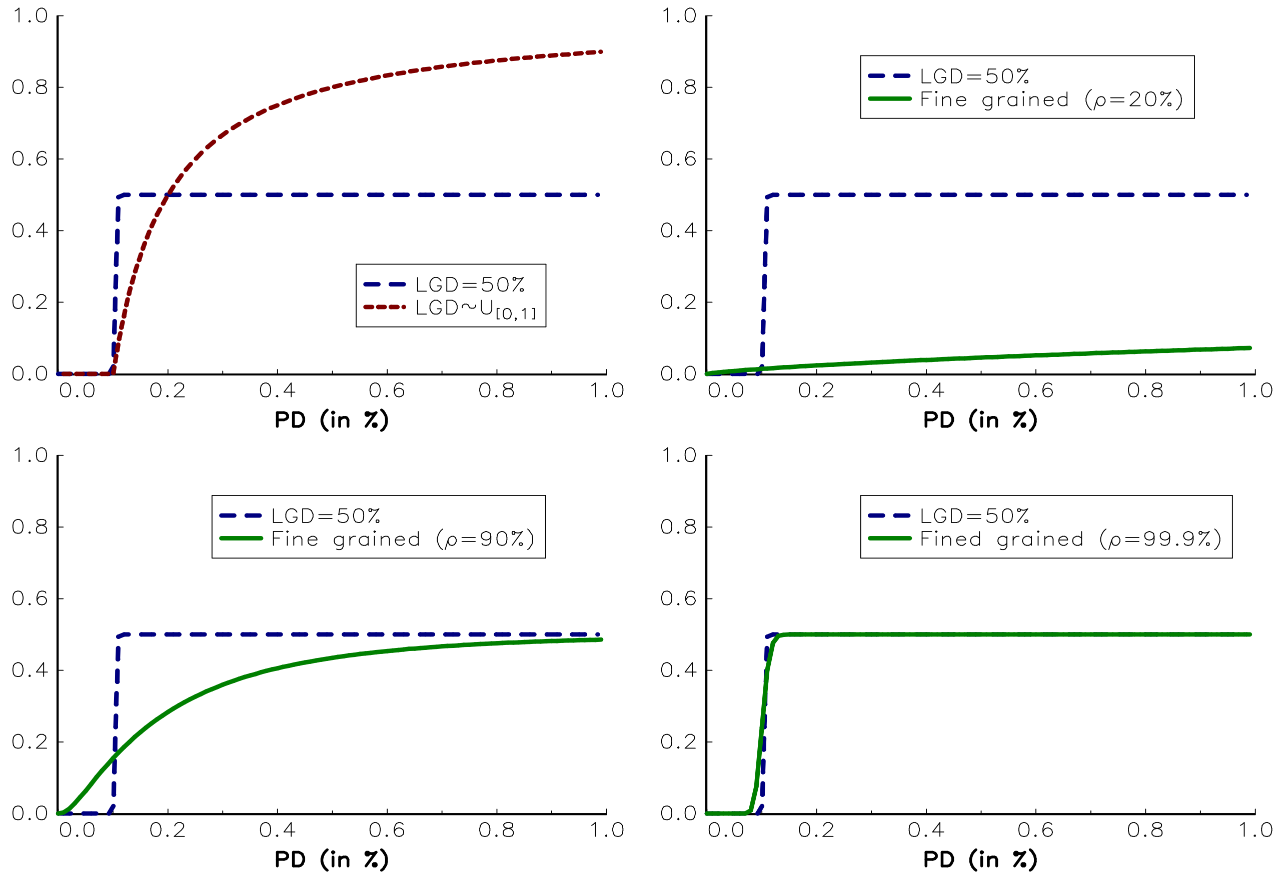


Figure: Comparison between the 99.9% value-at-risk of a loan and its risk contribution in an IFG portfolio

Granularity and concentration

IFG versus non-IFG portfolios

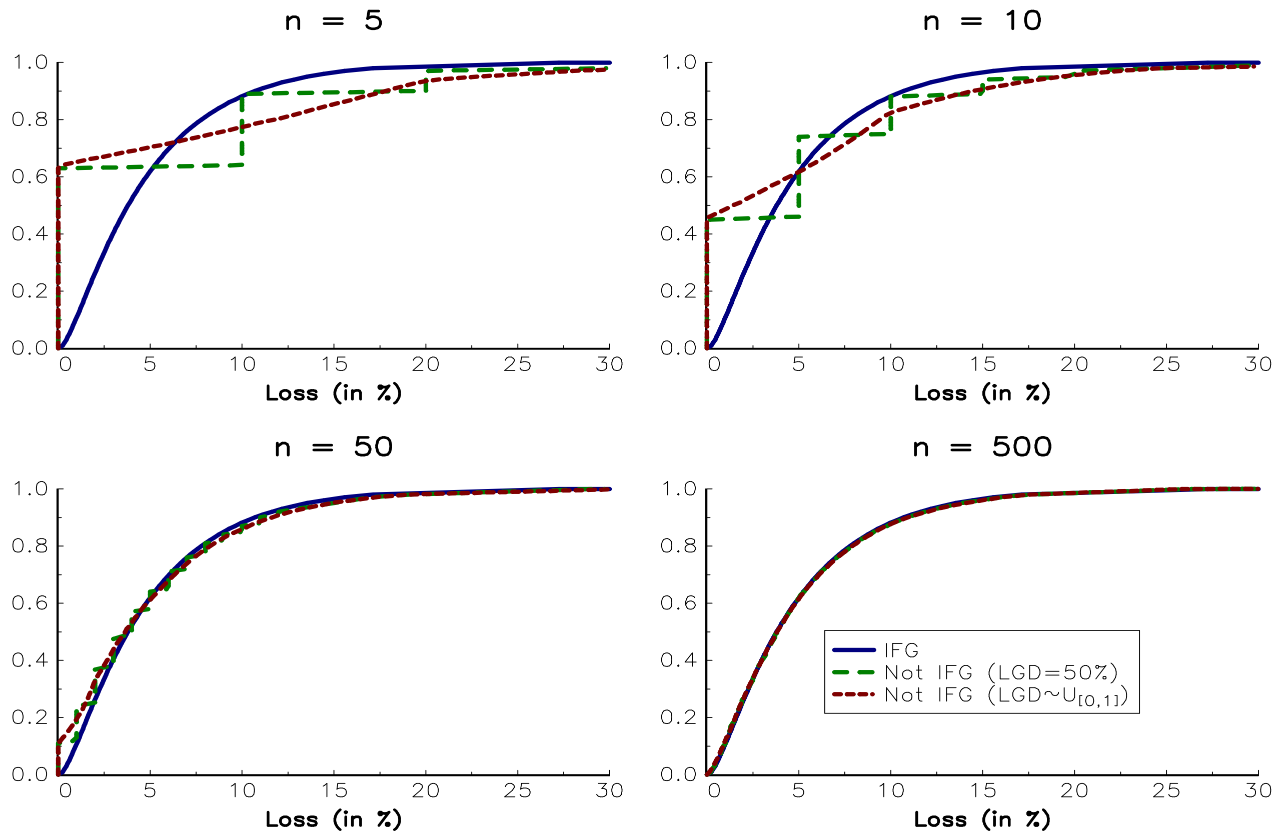


Figure: Comparison of the loss distribution of non-IFG and IFG portfolios

Exercises

- Credit derivatives
  - Exercise 3.4.1 – Single- and multi-name credit default swaps
- Basel II model
  - Exercise 3.4.8 – Variance of the conditional portfolio loss
  - Exercise 3.4.2 – Risk contribution in the Basel II model
  - Exercise 3.4.7 – Derivation of the original Basel granularity adjustement
- Parameter modeling
  - Exercise 3.4.3 – Calibration of the piecewise exponential model
  - Exercise 3.4.4 – Modeling loss given default
  - Exercise 3.4.5 – Modeling default times with a Markov chain
  - Exercise 3.4.6 – Continuous-time modeling of default risk

References

- Basel Committee on Banking Supervision (2001)
  The New Basel Capital Accord, Second consultative paper on Basel II, January 2001.

- Basel Committee on Banking Supervision (2004)
  International Convergence of Capital Measurement and Capital Standards — A Revised Framework, June 2004.

- Basel Committee on Banking Supervision (2015)
  Revisions to the Standardized Approach for Credit Risk, Consultative Document, December 2015.

- Basel Committee on Banking Supervision (2017)
  Basel III: Finalising Post-crisis Reforms, December 2017.

- Roncalli, T. (2020)
  Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 3.

Course 2023-2024 in Financial Risk Management
Lecture 4. Counterparty Credit Risk and Collateral Risk

Thierry Roncalli*

*Amundi Asset Management[13]

*University of Paris-Saclay

September 2023

[13] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- **Lecture 4: Counterparty Credit Risk and Collateral Risk**
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Counterparty credit risk and collateral risk are other forms of credit risk, where the underlying credit risk is not directly generated by the economic objective of the financial transaction

$\Rightarrow$ The portfolio can suffer a loss even if the business objective is reached

Some examples:

- 1997: LTCM (CCR)
- 2008: Lehman Brothers (CVA)
- 2011: ETF & Repo markets (Collateral risk)

Credit risk (CR) $\neq$ Counterparty credit risk (CCR)

CR:

- Loan $\Rightarrow$ credit risk (which is rewarded by a credit spread)
- CDS $\Rightarrow$ credit risk of the firm

CCR:

- Option $\Rightarrow$ counterparty credit risk (because the settlement is not guaranteed)
- CDS $\Rightarrow$ counterpart credit risk (if one counterparty defaults before the firm)

Definition

Definition

BCBS (2006) measures the counterparty credit risk by the replacement cost of the OTC derivative

Definition

Let us consider two banks A and B that have entered into an OTC contract $\mathfrak{C}$. We assume that the bank B defaults before the maturity of the contract. Bank A can then face two situations:

- The current value of the contract $\mathfrak{C}$ is negative $\Rightarrow$ Bank A closes out the position, pays the market value of the contract to Bank B, enters with another counterparty into a similar contract and receives the market value of the contract
- The current value of the contract $\mathfrak{C}$ is positive $\Rightarrow$ Bank A closes out the position, receives nothing from Bank B, enters with another counterparty into a similar contract and pays the market value of the contract

Loss = maximum between zero and the market value

This loss is not a market risk, a credit risk but a counterparty credit risk

CCR is more complex than CR

1. The counterparty credit risk is bilateral, meaning that both counterparties may face losses (Banks A and B)
2. The exposure at default is uncertain, because we don't know what will be the replacement cost of the contract when the counterparty defaults

The credit loss of an OTC portfolio is:

$$L = \sum_{i=1}^{n} \mathrm{EAD}_i \left(\boldsymbol{\tau}_i\right) \cdot \mathrm{LGD}_i \cdot \mathbb{1} \left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

$\Rightarrow$ The exposure at default is random and depends on different factors:

- The default time of the counterparty
- The evolution of market risk factors
- The correlation between the market value of the OTC contract and the default of the counterparty

Exposure at default

Exposure at default

We have:

$$\text{EAD} = \max\left(\text{MtM}\left(\boldsymbol{\tau}\right), 0\right)$$

Table: EAD of a portfolio

No netting	$\text{EAD} =$	$\sum_{i=1}^{n} \max\left(\text{MtM}_i\left(\boldsymbol{\tau}\right), 0\right)$
Global netting	$\text{EAD} =$	$\max\left(\sum_{i=1}^{n} \text{MtM}_i\left(\boldsymbol{\tau}\right), 0\right)$
Netting sets	$\text{EAD} =$	$\sum_{k} \max\left(\sum_{i\in\mathcal{N}_k} \text{MtM}_i\left(\boldsymbol{\tau}\right), 0\right) + \sum_{i\notin\cup\mathcal{N}_k} \max\left(\text{MtM}_i\left(\boldsymbol{\tau}\right), 0\right)$

Exposure at default

Example

Banks A and B have traded five OTC products, whose mark-to-market values[a] are given in the table below:

t	1	2	3	4	5	6	7	8
$\mathfrak{C}_1$	5	5	3	0	−4	0	5	8
$\mathfrak{C}_2$	−5	10	5	−3	−2	−8	−7	−10
$\mathfrak{C}_3$	0	2	−3	−4	−6	−3	0	5
$\mathfrak{C}_4$	2	−5	−5	−5	2	3	5	7
$\mathfrak{C}_5$	−1	−3	−4	−5	−7	−6	−7	−6

[a] They are calculated from the viewpoint of Bank A.

- No netting
- Global netting
- Partial netting = equity OTC contracts ($\mathfrak{C}_1$ and $\mathfrak{C}_2$) and fixed income OTC contracts ($\mathfrak{C}_3$ and $\mathfrak{C}_4$)

Exposure at default

Table: Counterparty exposure of Bank A

t	1	2	3	4	5	6	7	8
No netting	7	17	8	0	2	3	10	20
Global netting	1	9	0	0	0	0	0	4
Partial netting*	2	15	8	0	0	0	5	12

(*) Partial netting for $t = 8$: $\mathrm{EAD} = \max(8 - 10, 0) + \max(5 + 7, 0) + \max(-6, 0) = 12$

Table: Counterparty exposure of Bank B

t	1	2	3	4	5	6	7	8
No netting	6	8	12	17	19	17	14	16
Global netting	0	0	4	17	17	14	4	0
Partial netting	1	6	12	17	17	14	9	8

An illustrative example

Example

We consider a bank that buys 1 000 ATM call options, whose maturity is one-year. The current value of the underlying asset is equal to \$100. We assume that the interest rate r and the cost-of-carry parameter b are equal to 5%. Moreover, the implied volatility of the option is considered as a constant and is equal to 20%

We have:

$$\mathrm{MtM}(t) = n_C \cdot (\mathcal{C}(t) - \mathcal{C}_0)$$

where n_C and $\mathcal{C}(t)$ are the number and the market value of call options. The initial value of the call option is given by the Black-Scholes formula and we have $\mathcal{C}_0 = \$10.45$

The exposure at default $e(t)$ is equal to:

$$e(t) = \max(\mathrm{MtM}(t), 0)$$

An illustrative example

Table: Mark-to-market and counterparty exposure of the call option

t	Scenario #1				Scenario #2			
	$S(t)$	$\mathcal{C}(t)$	$\mathrm{MtM}(t)$	$e(t)$	$S(t)$	$\mathcal{C}(t)$	$\mathrm{MtM}(t)$	$e(t)$
1M	97.58	8.44	−2 013	0	91.63	5.36	−5 092	0
2M	98.19	8.25	−2 199	0	89.17	3.89	−6 564	0
3M	95.59	6.26	−4 188	0	97.60	7.35	−3 099	0
4M	106.97	12.97	2 519	2 519	97.59	6.77	−3 683	0
5M	104.95	10.83	382	382	96.29	5.48	−4 970	0
6M	110.73	14.68	4 232	4 232	97.14	5.29	−5 157	0
7M	113.20	16.15	5 700	5 700	107.71	11.55	1 098	1 098
8M	102.04	6.69	−3 761	0	105.71	9.27	−1 182	0
9M	115.76	17.25	6 802	6 802	107.87	10.18	−272	0
10M	103.58	5.96	−4 487	0	108.40	9.82	−630	0
11M	104.28	5.41	−5 043	0	104.68	5.73	−4 720	0
1Y	104.80	4.80	−5 646	0	115.46	15.46	5 013	5 013

An illustrative example

We have:

$$\mathrm{MtM}\left(0;t\right)=\mathrm{MtM}\left(0;t_0\right)+\mathrm{MtM}\left(t_0;t\right)$$

where 0 is the initial date of the trade, t_0 is the current date and t is the future date

$\Rightarrow$ This implies that the mark-to-market value at time t has two components:

1. The current mark-to-market value $\mathrm{MtM}\left(0;t_0\right)$ that depends on the past trajectory of the underlying price
2. The future mark-to-market value $\mathrm{MtM}\left(t_0;t\right)$ that depends on the future trajectory of the underlying price

How to calculate $\mathrm{MtM}\left(t_0;t\right)$?

- Historical probability measure $\mathbb{P}$
- Risk-neutral probability measure $\mathbb{Q}$

An illustrative example

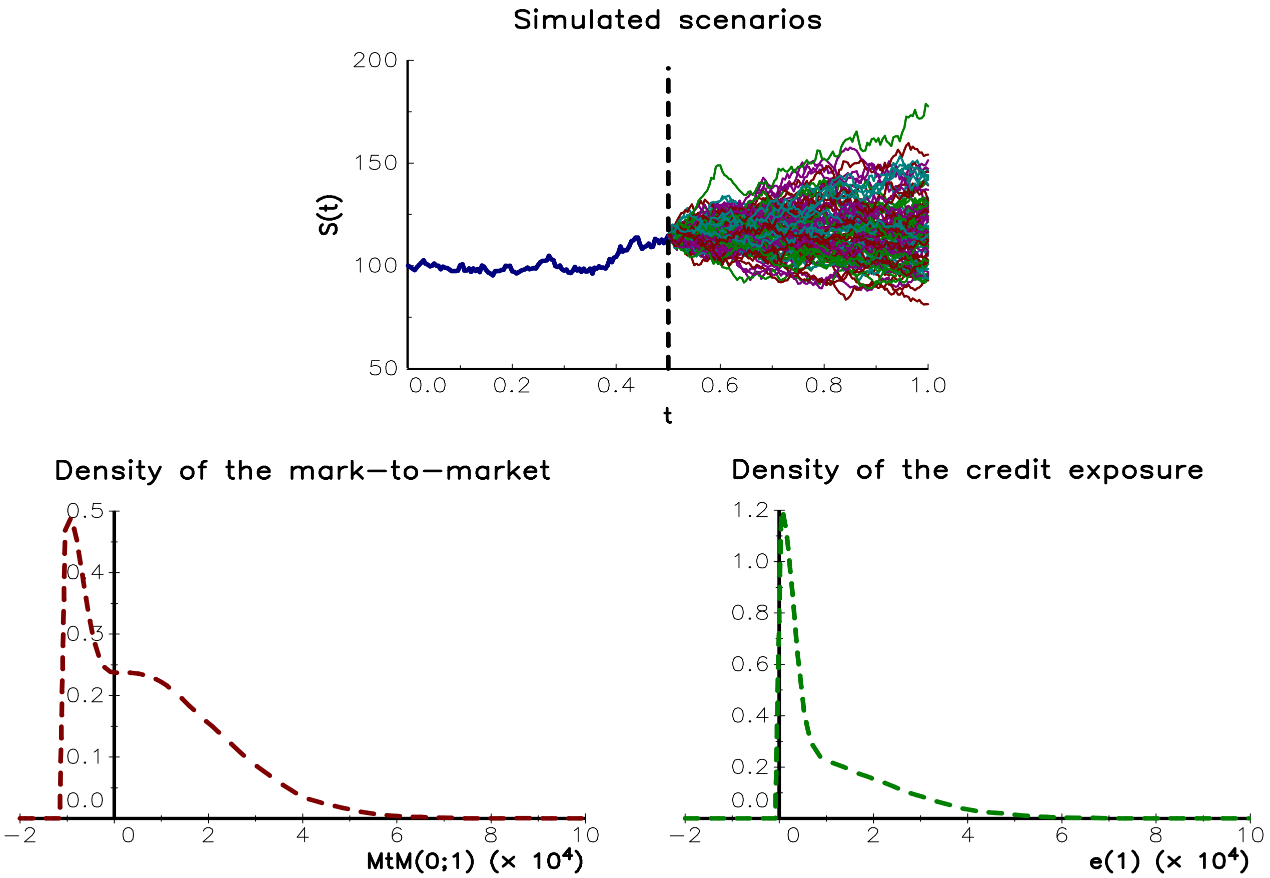


Figure: Probability density function of the counterparty exposure after six months

An illustrative example

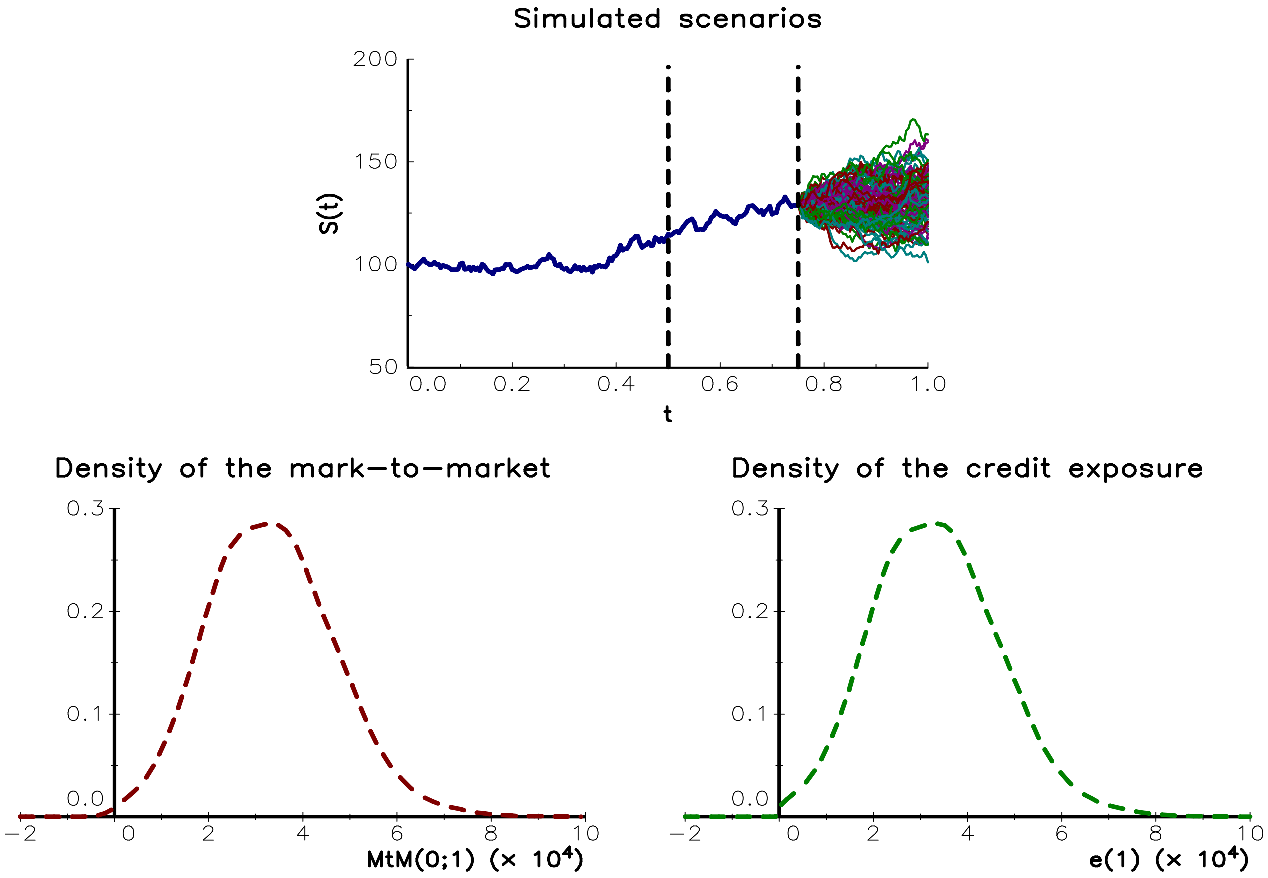


Figure: Probability density function of the counterparty exposure after nine months

An illustrative example

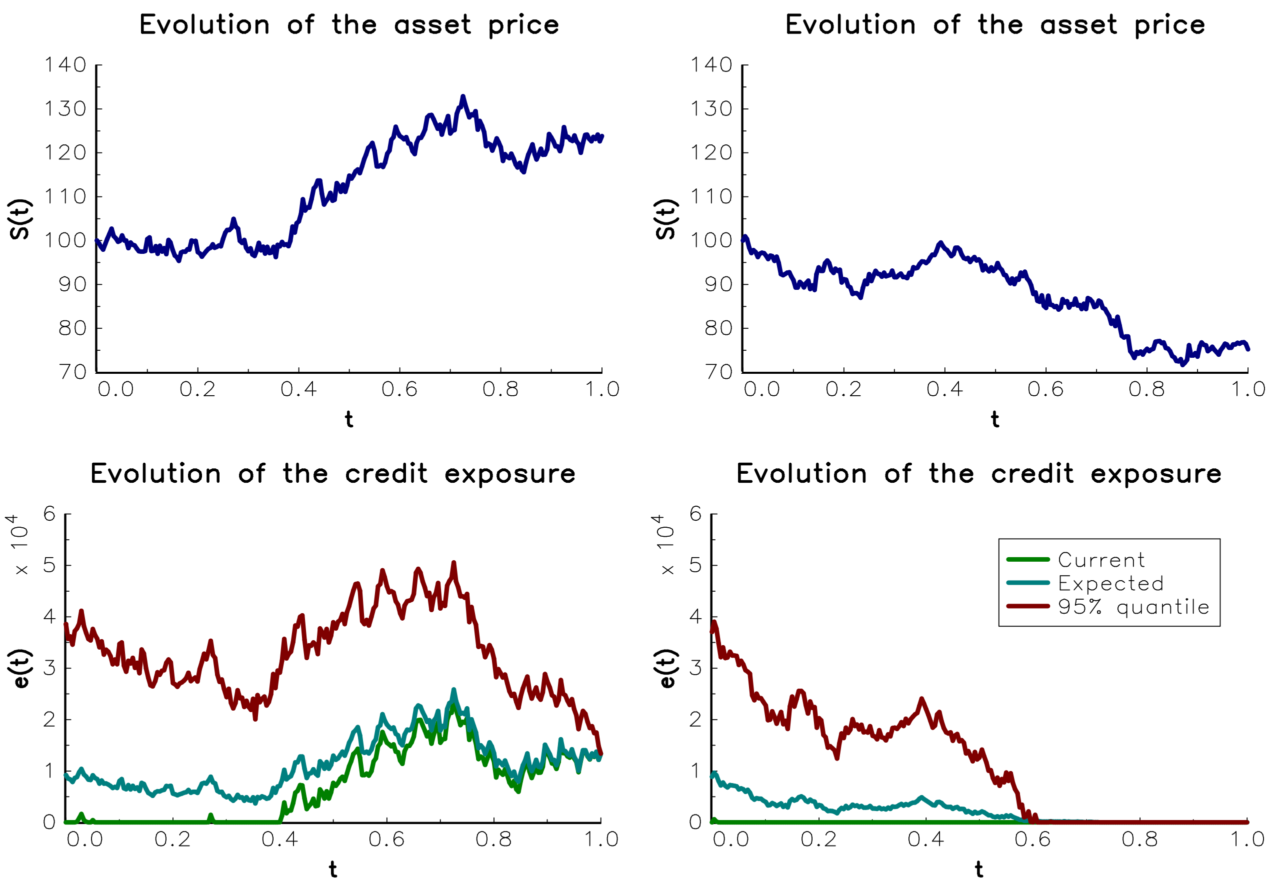


Figure: Evolution of the counterparty exposure

Measuring the counterparty exposure

- The counterparty exposure (or the potential future exposure – PFE) is equal to:

$$e(t) = \max\left(\mathrm{MtM}(0; t), 0\right)$$

- The current exposure is defined as:

$$\mathrm{CE}(t_0) = \max\left(\mathrm{MtM}(0; t_0), 0\right)$$

- $\mathbf{F}_{[0,t]}$ is the cumulative distribution function of the potential future exposure $e(t)$
- The peak exposure (PE) is the quantile of the counterparty exposure at the confidence level α:

$$\mathrm{PE}_\alpha(t) = \mathbf{F}_{[0,t]}^{-1}(\alpha) = \left\{\inf x : \Pr\left\{e(t) \leq x\right\} \geq \alpha\right\}$$

- The maximum peak exposure (MPE) is equal to:

$$\mathrm{MPE}_\alpha(0; t) = \sup_s \mathrm{PE}_\alpha(0; s)$$

Measuring the counterparty exposure

- The expected exposure (EE) is the average of the distribution of the counterparty exposure at the future date t:

$$\mathrm{EE}(t) = \mathbb{E}[e(t)] = \int_0^\infty x \,\mathrm{d}\mathbf{F}_{[0,t]}(x)$$

- The expected positive exposure (EPE) is the weighted average over time $[0, t]$ of the expected exposure:

$$\mathrm{EPE}(0;t) = \mathbb{E}\left[\frac{1}{t}\int_0^t e(s)\,\mathrm{d}s\right] = \frac{1}{t}\int_0^t \mathrm{EE}(s)\,\mathrm{d}s$$

- The effective expected exposure (EEE) is the maximum expected exposure that occurs at the future date t or any prior date:

$$\mathrm{EEE}(t) = \sup_{s\leq t}\mathrm{EE}(s) = \max\left(\mathrm{EEE}\left(t^-\right), \mathrm{EE}(t)\right)$$

- The effective expected positive exposure (EEPE) is the weighted average over time $[0, t]$ of the effective expected exposure:

$$\mathrm{EEPE}(0;t) = \frac{1}{t}\int_0^t \mathrm{EEE}(s)\,\mathrm{d}s$$

Exercise I

Exercise (HFRM, Exercise 4.4.2, Question 3, page 301)

We assume that:

$$e(t) = \exp\left(\sigma \cdot \sqrt{t} \cdot X\right)$$

where $X \sim \mathcal{N}(0,1)$

Solution of $\mathbf{F}_{[0,t]}$

- We have:

$$\begin{aligned}\mathbf{F}_{[0,t]}(x) &= \Pr\left\{e^{\sigma\sqrt{t}X} \leq x\right\} \\ &= \Pr\left\{\sigma\sqrt{t}X \leq \ln x\right\} \\ &= \Phi\left(\frac{\ln x}{\sigma\sqrt{t}}\right)\end{aligned}$$

  with $x \in [0, \infty]$

- We deduce that the probability density function is equal to:

$$\begin{aligned}f_{[0,t]}(x) &= \frac{\partial\,\mathbf{F}_{[0,t]}(x)}{\partial\,x} \\ &= \frac{1}{x\sigma\sqrt{t}}\phi\left(\frac{\ln x}{\sigma\sqrt{t}}\right)\end{aligned}$$

  We recognize the pdf of the log-normal distribution:

$$e(t) \sim \mathcal{LN}\left(0, \sigma^2 t\right)$$

Solution of PE

- We have:

$$\mathrm{PE}_{\alpha}(t) = \mathbf{F}_{[0,t]}^{-1}(\alpha)$$

It follows that:

$$\begin{aligned}\Phi\left(\frac{\ln x}{\sigma\sqrt{t}}\right) = \alpha & \Leftrightarrow \frac{\ln x}{\sigma\sqrt{t}} = \Phi^{-1}(\alpha) \\ & \Leftrightarrow x = \exp\left(\Phi^{-1}(\alpha)\,\sigma\sqrt{t}\right)\end{aligned}$$

We conclude that:

$$\mathrm{PE}_{\alpha}(t) = e^{\Phi^{-1}(\alpha)\sigma\sqrt{t}}$$

- It is obvious that $e^{\Phi^{-1}(\alpha)\sigma\sqrt{t}}$ is maximum when t is equal to the maturity T:

$$\mathrm{MPE}_{\alpha}(0;T) = \sup_t \mathrm{PE}_{\alpha}(t) = e^{\Phi^{-1}(\alpha)\sigma\sqrt{T}}$$

Solution of EE

- The expected exposure is the average of the potential future exposure:

$$\begin{aligned}\mathrm{EE}(t) &= \mathbb{E}[e(t)] \\ &= \int x \,\mathrm{d}\mathbf{F}_{[0,t]}(x) \\ &= \int x \, f_{[0,t]}(x) \,\mathrm{d}x\end{aligned}$$

We can compute the integral or we can use the property that $e(t) \sim \mathcal{LN}\left(0, \sigma^2 t\right)$. Since we know that:

$$\mathbb{E}\left[\mathcal{LN}\left(\mu, \sigma^2\right)\right] = \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

we conclude that:

$$\mathrm{EE}(t) = \exp\left(\frac{1}{2}\sigma^2 t\right)$$

Solution of EPE

- We have:

$$
\begin{aligned}
\mathrm{EPE}(0;t) &= \frac{1}{t}\int_0^t \mathrm{EE}(s)\,\mathrm{d}s \\
&= \frac{1}{t}\int_0^t e^{\frac{1}{2}\sigma^2 s}\,\mathrm{d}s \\
&= \frac{1}{t}\left[\frac{e^{\frac{1}{2}\sigma^2 s}}{\frac{1}{2}\sigma^2}\right]_0^t \\
&= \frac{2e^{\frac{1}{2}\sigma^2 t}-2}{\sigma^2 t}
\end{aligned}
$$

Solution of EEE

- Since the function $e^{\frac{1}{2}\sigma^2 t}$ is increasing with respect to t, we deduce that the effective expected exposure is equal to the expected exposure:

$$\begin{aligned}\mathrm{EEE}(t) &= \sup_{s \leq t} \mathrm{EE}(s) \\ &= \exp\left(\frac{1}{2}\sigma^2 t\right)\end{aligned}$$

Solution of EEPE

- It follows that:

$$
\begin{aligned}
\mathrm{EEPE}(0;t) &= \frac{1}{t}\int_0^t \mathrm{EEE}(s)\,\mathrm{d}s \\
&= \frac{1}{t}\int_0^t \mathrm{EE}(s)\,\mathrm{d}s \\
&= \mathrm{EPE}(0;t) \\
&= \frac{2e^{\frac{1}{2}\sigma^2 t}-2}{\sigma^2 t}
\end{aligned}
$$

Solution

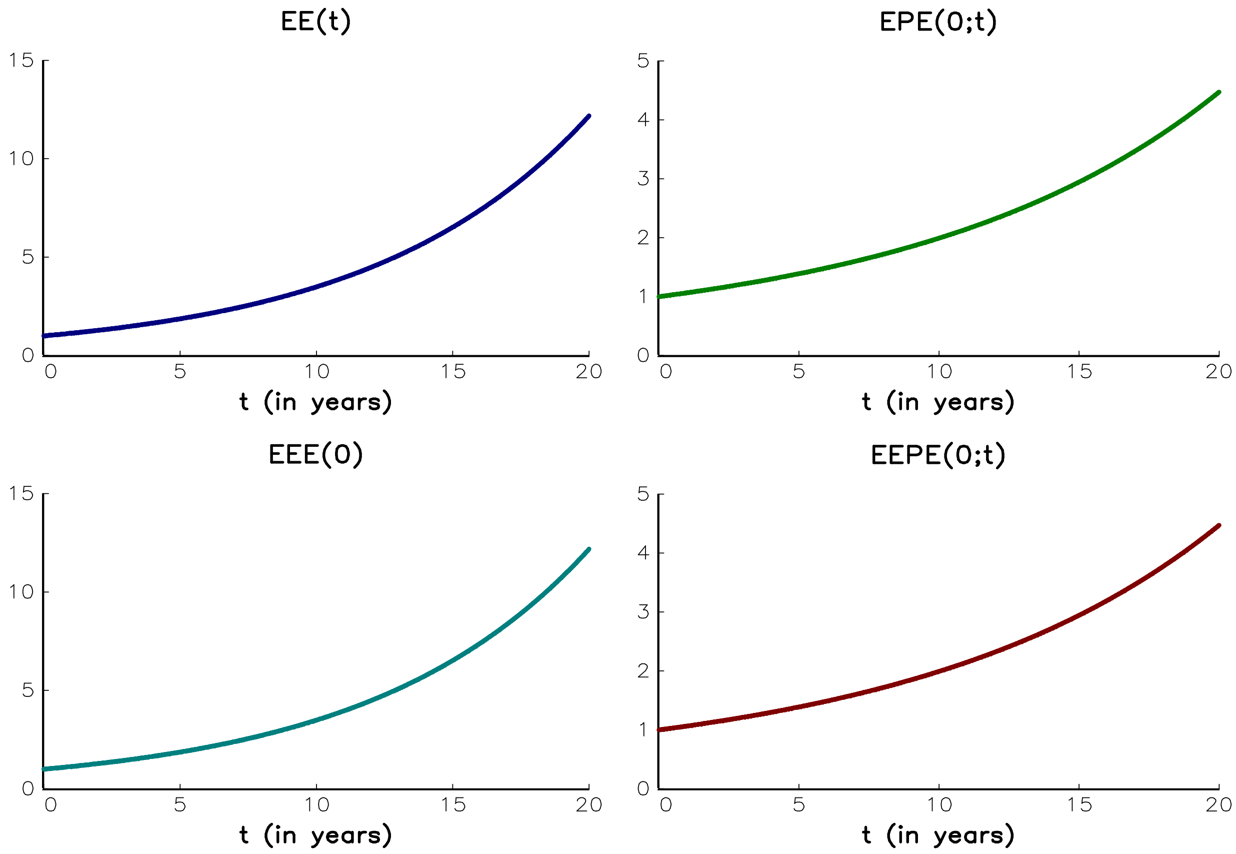


Figure: Credit exposure when $e(t) = \exp\left(\sigma\sqrt{t}\mathcal{N}(0,1)\right)$

Exercise II

Exercise (HFRM, Exercise 4.4.2, Question 4, page 301)

We assume that:

$$e(t) = \sigma \cdot \left(t^3 - \frac{7}{3} T t^2 + \frac{4}{3} T^2 t \right) \cdot X$$

where $X \sim \mathcal{U}_{[0,1]}$

Solution

Solution (HFRM-CB, pages 75-76)

$$\mathbf{F}_{[0,t]}(x) = \frac{x}{\sigma\left(t^3 - \frac{7}{3}Tt^2 + \frac{4}{3}T^2 t\right)} \text{ with } x \in \left[0, \sigma\left(t^3 - \frac{7}{3}Tt^2 + \frac{4}{3}T^2 t\right)\right]$$

$$\mathrm{PE}_\alpha(0) = \alpha\sigma\left(t^3 - \frac{7}{3}Tt^2 + \frac{4}{3}T^2 t\right)$$

$$\mathrm{MPE}_\alpha(0;t) = \mathbb{1}\{t < t^\star\} \times \mathrm{PFE}_\alpha(0;t) + \mathbb{1}\{t \geq t^\star\} \times \mathrm{PFE}_\alpha(0;t^\star)$$

$$\mathrm{EE}(t) = \frac{1}{2}\sigma\left(t^3 - \frac{7}{3}Tt^2 + \frac{4}{3}T^2 t\right)$$

$$\mathrm{EPE}(0;t) = \sigma\left(\frac{9t^3 - 28Tt^2 + 24T^2 t}{72}\right)$$

$$\mathrm{EEE}(t) = \mathbb{1}\{t < t^\star\} \times \mathrm{EE}(t) + \mathbb{1}\{t \geq t^\star\} \times \mathrm{EE}(t^\star)$$

$$t^\star = \left(\frac{7 - \sqrt{13}}{9}\right) T$$

Solution

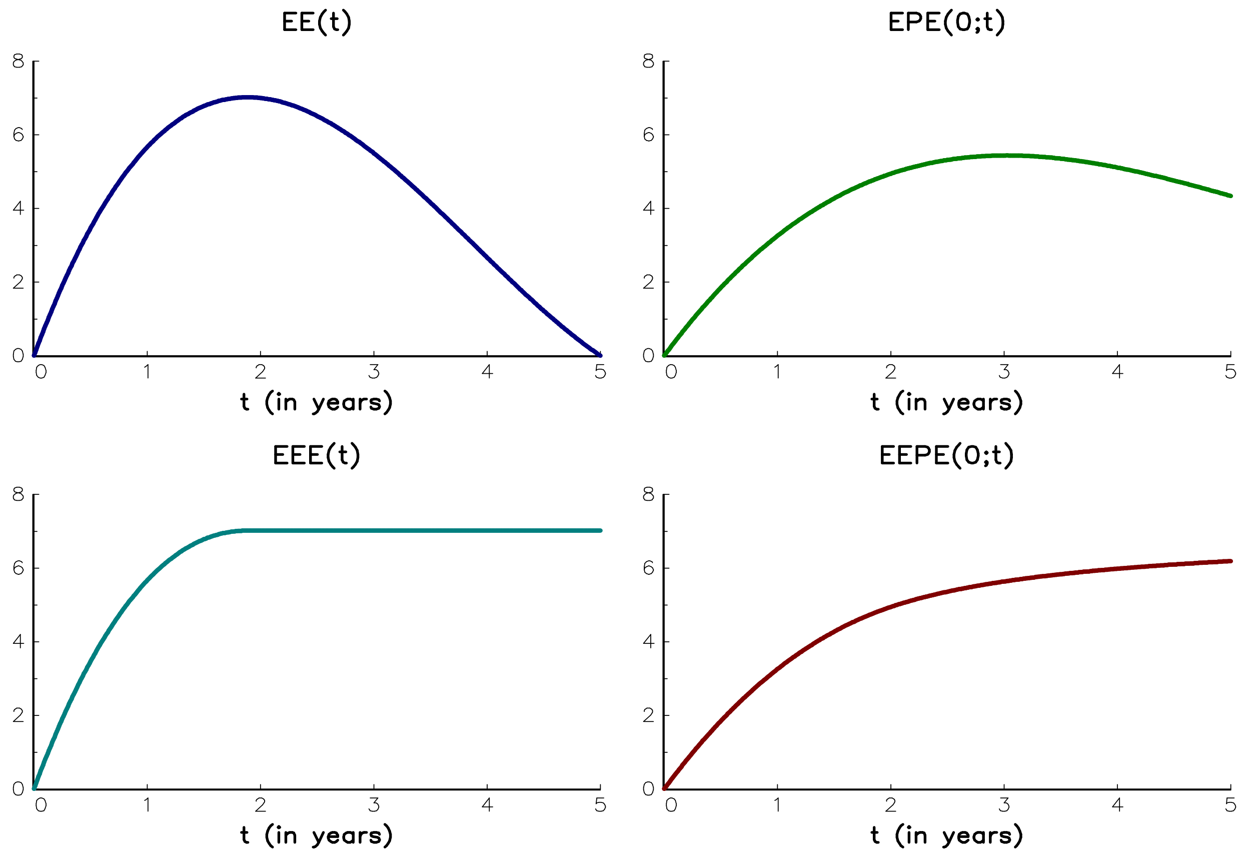


Figure: Credit exposure when $e(t) = \sigma \left(t^3 - \frac{7}{3} T t^2 + \frac{4}{3} T^2 t \right) \mathcal{U}_{[0,1]}$

Practical implementation for calculating counterparty exposures

- In practice, we use Monte Carlo simulations and the risk-neutral distribution probability $\mathbb{Q}$
- We consider a set of discrete times $\{t_0, t_1, \ldots, t_n\}$
- We note $\mathrm{MtM}_j(t_i)$ the simulated mark-to-market value for the j^{th} simulation at time at time t_i
- We note n_S the number of Monte Carlo simulations

Remark

If we consider the introductory example, we simulate $S_j(t_i)$ the value of the asset price at time t_i for the j^{th} simulation. For each simulated trajectory, we then calculate the option price $\mathcal{C}_j(t_i)$ and the mark-to-market value:

$$\mathrm{MtM}_j(t_i) = n_C \cdot (\mathcal{C}_j(t_i) - \mathcal{C}_0)$$

Practical implementation

Given a sample of n_S simulated exposures for $t \in \{t_0, t_1, \ldots, t_n\}$:

$$e_j(t_i) = \max(\mathrm{MtM}_j(t_i), 0)$$

we deduce the following estimators:

- The peak exposure at time t_i is estimated using the order statistics:

$$\mathrm{PE}_\alpha(t_i) = e_{\alpha n_S : n_S}(t_i)$$

- We use the empirical mean to calculate the expected exposure:

$$\mathrm{EE}(t_i) = \frac{1}{n_S} \sum_{j=1}^{n_S} e_j(t_i)$$

Practical implementation

- For the expected positive exposure, we approximate the integral by the following sum:

$$\mathrm{EPE}\left(0;t_i\right)=\frac{1}{t_i}\sum_{k=1}^{i}\mathrm{EE}\left(t_k\right)\Delta t_k$$

If we consider a fixed-interval scheme with $\Delta t_k=\Delta t$, we obtain:

$$\mathrm{EPE}\left(0;t_i\right)=\frac{\Delta t}{t_i}\sum_{k=1}^{i}\mathrm{EE}\left(t_k\right)=\frac{1}{i}\sum_{k=1}^{i}\mathrm{EE}\left(t_k\right)$$

Practical implementation

- By definition, the effective expected exposure is given by the following recursive formula:

$$\mathrm{EEE}(t_i) = \max\left(\mathrm{EEE}(t_{i-1}), \mathrm{EE}(t_i)\right)$$

  where $\mathrm{EEE}(0)$ is initialized with the value $\mathrm{EE}(0)$
- Finally, the effective expected positive exposure is given by:

$$\mathrm{EEPE}(0; t_i) = \frac{1}{t_i} \sum_{k=1}^{i} \mathrm{EEE}(t_k) \Delta t_k$$

  In the case of a fixed-interval scheme, this formula becomes:

$$\mathrm{EEPE}(0; t_i) = \frac{1}{i} \sum_{k=1}^{i} \mathrm{EEE}(t_k)$$

The square-root profile of CCR

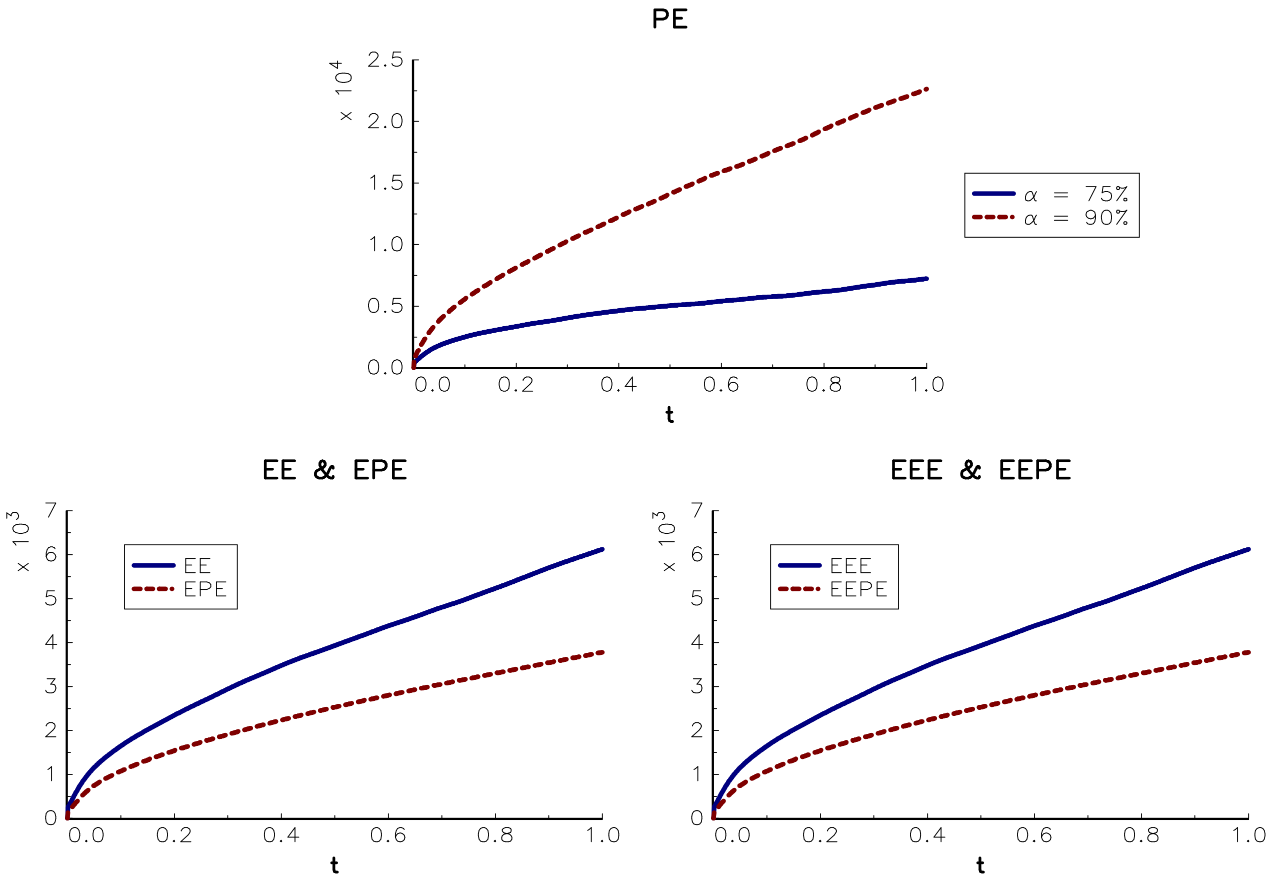


Figure: Counterparty exposure profile of options

The bell-shaped profile of CCR

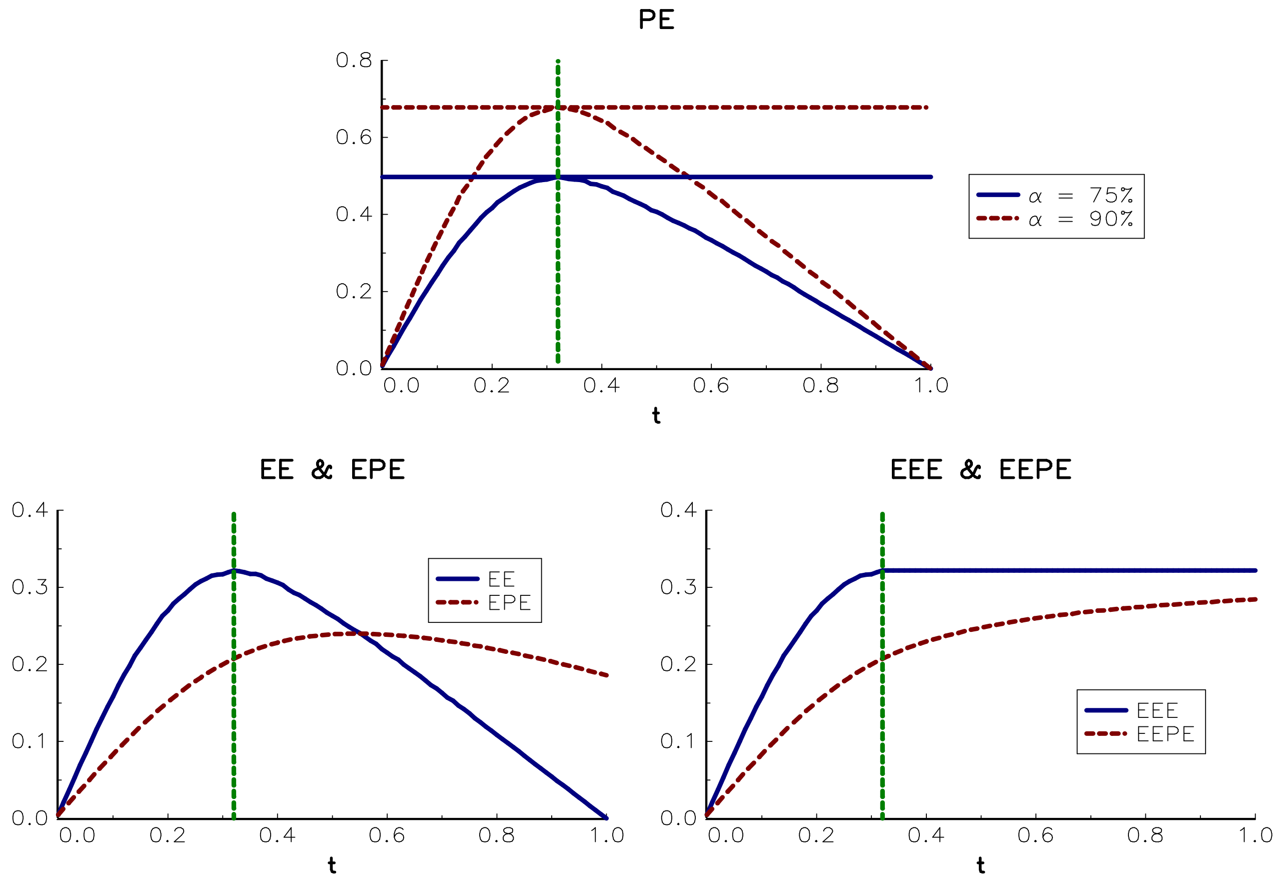


Figure: Counterparty exposure profile of interest rate swaps

Regulatory capital

Basel II

- Non-internal model methods
  1. Current exposure method (CEM)
  2. Standardized method (SM)
- Internal model method (IMM)

Basel III

- Standardized approach (SA-CCR)
- Internal model method (IMM)

Each approach defines how the exposure at default EAD is calculated. In the SA approach, the capital charge is equal to:

$$\mathcal{K} = 8\% \cdot \mathrm{EAD} \cdot \mathrm{RW}$$

In the IRB approach, we recall that:

$$\mathcal{K} = \mathrm{EAD} \cdot \mathrm{LGD} \cdot \left(\Phi \left(\frac{\Phi^{-1}(\mathrm{PD}) + \sqrt{\rho(\mathrm{PD})}\Phi^{-1}(0.999)}{\sqrt{1 - \rho(\mathrm{PD})}} \right) - \mathrm{PD} \right) \cdot \varphi(\mathrm{M})$$

Regulatory capital
Internal model method (Basel II and III)

We have:

$$\mathrm{EAD} = \alpha \cdot \mathrm{EEPE}\left(0; \min\left(T, 1\right)\right)$$

where α is equal to 1.4 and T is the maturity of the OTC contract

Remark

Under some conditions, the bank may uses its own estimates for α, but it must be larger than 1.2

Regulatory capital

Internal model method (Basel II and III)

Example

We assume that the one-year effective expected positive exposure with respect to a given counterparty is equal to \$50.2 mn. The LGD is equal to 45% and the maturity is set to one year.

Table: Capital charge of counterparty credit risk under the FIRB approach

	PD	1%	2%	3%	4%	5%
Basel II	ρ (PD) (in %)	19.28	16.41	14.68	13.62	12.99
	$\mathcal{K}$ (in \$ mn)	4.12	5.38	6.18	6.82	7.42
Basel III	ρ (PD) (in %)	24.10	20.52	18.35	17.03	16.23
	$\mathcal{K}$ (in \$ mn)	5.26	6.69	7.55	8.25	8.89
	$\Delta\mathcal{K}$ (in %)	27.77	24.29	22.26	20.89	19.88

Regulatory capital

SA-CCR method (Basel III)

The exposure at default under the SA-CCR is defined as follows:

$$\mathrm{EAD} = \alpha \cdot (\mathrm{RC} + \mathrm{PFE})$$

where RC is the replacement cost (or the current exposure), PFE is the potential future exposure and α is equal to 1.4

Remark

We can view this formula as an approximation of the IMM calculation, meaning that $\mathrm{RC} + \mathrm{PFE}$ *represents a stylized* EEPE *value*

$\Rightarrow$ SA-CCR is close to SA-TB (see HFRM on pages 270-274)

Impact of wrong way risk

Definition

The wrong way risk (WWR) is defined as the risk that "*occurs when exposure to a counterparty or collateral associated with a transaction is adversely correlated with the credit quality of that counterparty*". This means that the exposure at default of the OTC contract and the default risk of the counterparty are positively correlated

Two types of wrong way risk:

1. General (or conjectural) wrong way risk occurs when the credit quality of the counterparty is correlated with macroeconomic factors, which also impact the value of the transaction (e.g. level of interest rates)
2. Specific wrong way risk occurs when the correlation between the exposure at default and the probability of default is mainly explained by some idiosyncratic factors (e.g. Bank A buys a CDS protection on Bank B from Bank C)

Impact of wrong way risk

An example

We assume that:

$$\mathrm{MtM}(t) = \mu + \sigma W(t)$$

If we note $e(t) = \max(\mathrm{MtM}(t), 0)$, we have:

$$\begin{aligned}
\mathbb{E}[e(t)] &= \int_{-\infty}^{\infty} \max\left(\mu + \sigma\sqrt{t}x, 0\right)\phi(x)\,\mathrm{d}x \\
&= \mu \int_{-\mu/(\sigma\sqrt{t})}^{\infty} \phi(x)\,\mathrm{d}x + \sigma\sqrt{t}\int_{-\mu/(\sigma\sqrt{t})}^{\infty} x\phi(x)\,\mathrm{d}x \\
&= \mu\left(1 - \Phi\left(-\frac{\mu}{\sigma\sqrt{t}}\right)\right) + \sigma\sqrt{t}\left[-\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\right]_{-\mu/(\sigma\sqrt{t})}^{\infty} \\
&= \mu\Phi\left(\frac{\mu}{\sigma\sqrt{t}}\right) + \sigma\sqrt{t}\phi\left(\frac{\mu}{\sigma\sqrt{t}}\right)
\end{aligned}$$

Impact of wrong way risk

An example

Two assumptions:

$\mathcal{H}_1$ Merton model with the default barrier $B(t) = \Phi^{-1}(1 - \mathbf{S}(t))$

$\mathcal{H}_2$ The dependence between the mark-to-market $\mathrm{MtM}(t)$ and the survival time is given by the Normal copula $\mathbf{C}(u_1, u_2; \rho)$ with parameter ρ

Impact of wrong way risk

An example

Since we have $1 - \mathbf{S}(t) \sim \mathcal{U}_{[0,1]}$, it follows that $B(t) \sim \mathcal{N}(0,1)$. We deduce that the random vector $(\mathrm{MtM}(t), B(t))$ is normally distributed:

$$\begin{pmatrix} \mathrm{MtM}(t) \\ B(t) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 t & \rho\sigma\sqrt{t} \\ \rho\sigma\sqrt{t} & 1 \end{pmatrix} \right)$$

because the correlation $\rho(\mathrm{MtM}(t), B(t))$ is equal to the Normal copula parameter ρ. Using the conditional expectation formula (Lecture 2, Slide 114), it follows that:

$$\mathrm{MtM}(t) \mid B(t) = B \sim \mathcal{N}\left(\mu_B, \sigma_B^2\right)$$

where:

$$\mu_B = \mu + \rho\sigma\sqrt{t}(B - 0)$$

and:

$$\sigma_B^2 = \sigma^2 t - \rho^2\sigma^2 t = \left(1 - \rho^2\right)\sigma^2 t$$

Impact of wrong way risk

An example

We deduce that:

$$\mathbb{E}\left[e\left(t\right) \mid \boldsymbol{\tau} = t\right] = \mathbb{E}\left[e\left(t\right) \mid B\left(t\right) = B\right] = \mu_B \Phi\left(\frac{\mu_B}{\sigma_B}\right) + \sigma_B \phi\left(\frac{\mu_B}{\sigma_B}\right)$$

where:

$$\mu_B = \mu + \rho\sigma\sqrt{t}B$$

and:

$$\sigma_B = \sqrt{1-\rho^2}\sigma\sqrt{t}$$

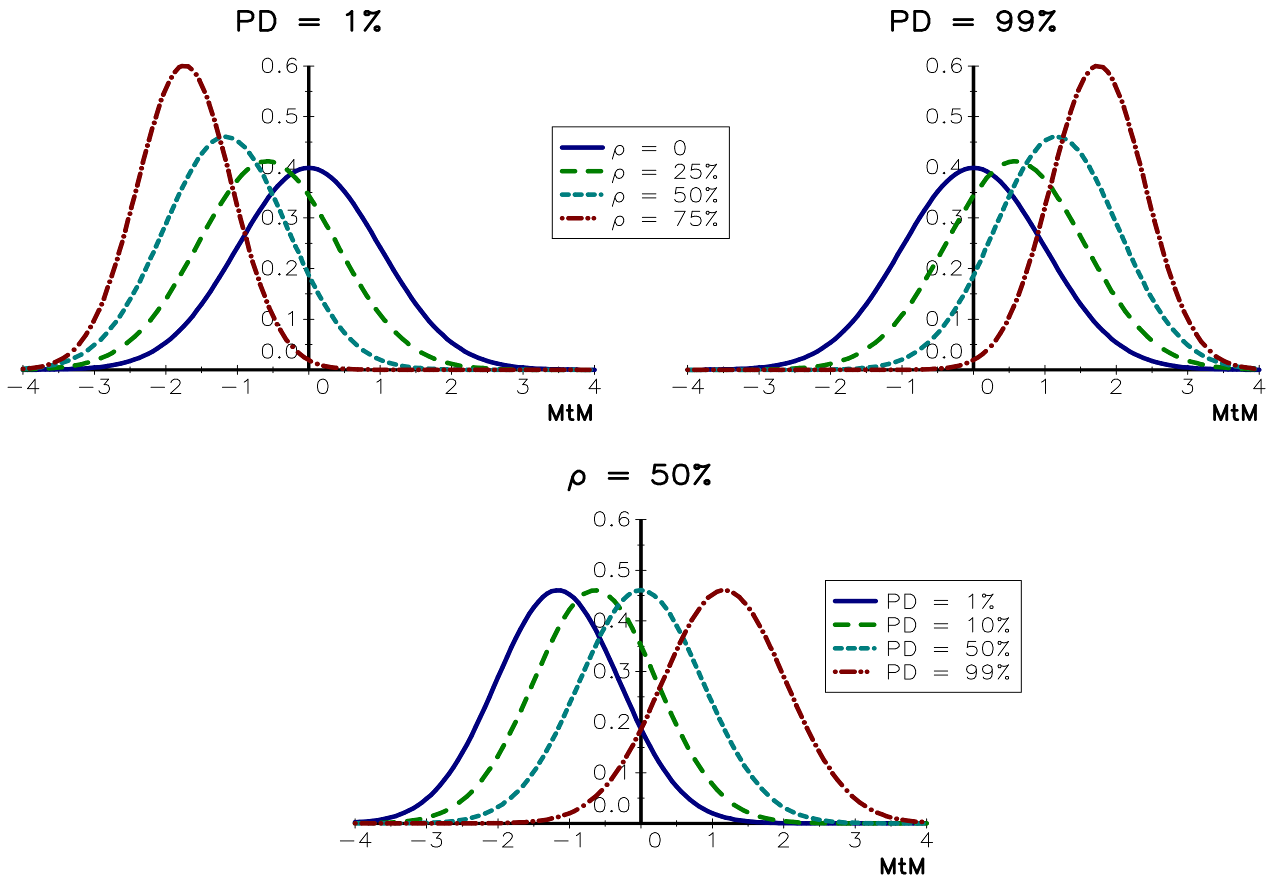
With the exception of $\rho = 0$, we have:

$$\mathbb{E}\left[e\left(t\right)\right] \neq \mathbb{E}\left[e\left(t\right) \mid \boldsymbol{\tau} = t\right]$$

Impact of wrong way risk

An example

Figure: Conditional distribution of the mark-to-market

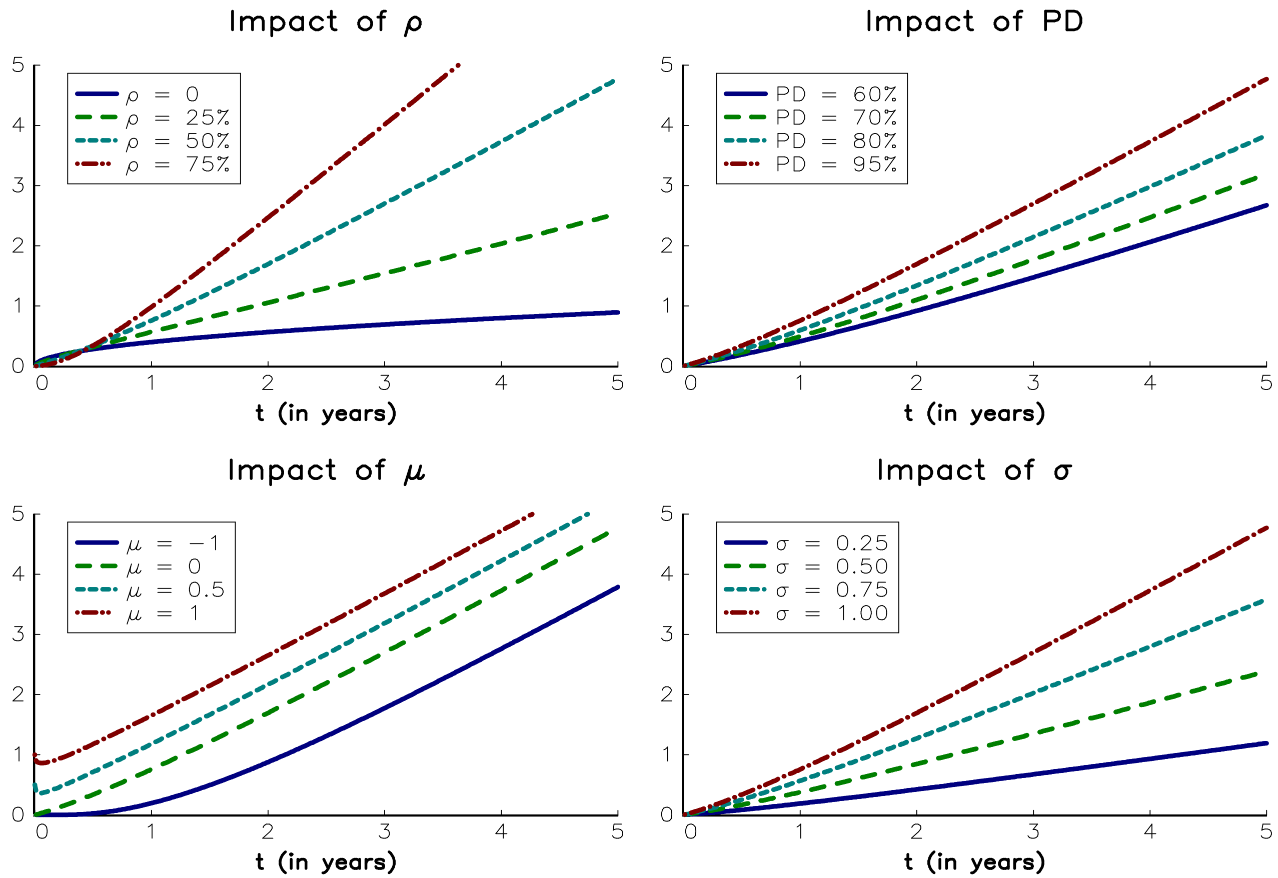


(*) The default occurs at time $t = 1$, and the parameters are $\mu = 0$, $\sigma = 1$ and $\boldsymbol{\tau} \sim \mathcal{E}(\lambda)$

Impact of wrong way risk

An example

Figure: Conditional expectation of the exposure at default



(*) The default values are $\mu = 0$, $\sigma = 1$, $\mathrm{PD} = 90\%$ and $\rho = 50\%$

Impact of wrong way risk

Calibration of the α factor

$\Rightarrow$ A difficult task:

$$L = \sum_{i=1}^{n} \mathrm{EAD}\left(\boldsymbol{\tau}_i, \mathcal{F}_1, \ldots, \mathcal{F}_m\right) \cdot \mathrm{LGD}_i \cdot \mathbb{1}\left\{\boldsymbol{\tau}_i \leq T_i\right\}$$

where $\mathcal{F} = \left(\mathcal{F}_1, \ldots, \mathcal{F}_m\right)$ are the market risk factors and $\boldsymbol{\tau} = \left(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_n\right)$ are the default times

WWR implies to correlate the random vectors $\mathcal{F}$ and $\boldsymbol{\tau}$

CVA versus CCR

Definition

CVA is the adjustment to the risk-free (or fair) value of derivative instruments to account for counterparty credit risk. Thus, CVA is commonly viewed as the market price of CCR

- CCR concerns the default risk of the counterparty ⇒ credit risk

  CCR may induce a loss

- CVA concerns the credit risk of the counterparty before the default ⇒ market risk

  CVA impacts the mark-to-market of the OTC contract

2008 GFC & Lehman Brothers bankruptcy

Banks suffered significant CCR losses on their OTC derivatives portfolios:

- $^{2}/_{3}$ of these losses came from CVA markdowns on derivatives
- $^{1}/_{3}$ were due to counterparty defaults

Fair valuation

We consider two banks A and B and an OTC contract $\mathfrak{C}$. The P&L $\Pi_{A|B}$ of Bank A is equal to:

$$\Pi_{A|B} = \mathrm{MtM} - \mathrm{CVA}_B$$

where MtM is the risk-free mark-to-market value of $\mathfrak{C}$ and CVA_B is the CVA with respect to Bank B. We assume that Bank A has traded the same contract with Bank C. It follows that:

$$\Pi_{A|C} = \mathrm{MtM} - \mathrm{CVA}_C$$

In a world where there is no counterparty credit risk, we have:

$$\Pi_{A|B} = \Pi_{A|C} = \mathrm{MtM}$$

Fair valuation

If we take into account the counterparty credit risk, the two P&Ls of the same contract are different because Bank A does not face the same risk:

$$\Pi_{A|B} \neq \Pi_{A|C}$$

In particular, if Bank A wants to close the two exposures, it is obvious that the contact $\mathfrak{C}$ with the counterparty B has more value than the contact $\mathfrak{C}$ with the counterparty C if the credit risk of B is lower than the credit risk of C

CVA, DVA and bilateral CVA

- CVA is the market risk related to the credit risk of the counterparty
- DVA (debit value adjustment) is the credit-related adjustment capturing the entity's own credit risk
- BCVA (bilateral CVA) is the combination of the two credit-related adjustments:

$$\Pi_{A|B} = \mathrm{MtM} + \underbrace{\mathrm{DVA}_A - \mathrm{CVA}_B}_{\text{Bilateral CVA}}$$

- If the credit risk of Bank A is lower than the credit risk of Bank B, the bilateral CVA of Bank A is negative and reduces the value of the OTC portfolio from the perspective of Bank A
- If the credit risk of Bank A is higher than the credit risk of Bank B, the bilateral CVA of Bank A is positive and increases the value of the OTC portfolio from the perspective of Bank A
- If the credit risk of Banks A and B is the same, the bilateral CVA is equal to zero

BCVA and the coherency property

The DVA of Bank A is the CVA of Bank A from the perspective of Bank B:

$$\mathrm{CVA}_A = \mathrm{DVA}_A$$

We also have $\mathrm{DVA}_B = \mathrm{CVA}_B$, which implies that the P&L of Bank B is equal to:

$$\begin{aligned}
\Pi_{B|A} &= -\mathrm{MtM} + \mathrm{DVA}_B - \mathrm{CVA}_A \\
&= -\mathrm{MtM} + \mathrm{CVA}_B - \mathrm{DVA}_A \\
&= -\Pi_{A|B}
\end{aligned}$$

Remark

We deduce that the P&Ls of Banks A and B are coherent in the bilateral CVA framework as in the risk-free MtM framework

Notations

- The positive exposure $e^{+}(t)$ is the maximum between 0 and the risk-free mark-to-market:

$$e^{+}(t) = \max(\mathrm{MtM}(t), 0)$$

  This quantity was previously denoted by $e(t)$ and corresponds to the potential future exposure in the CCR framework

- The negative exposure $e^{-}(t)$ is the difference between the risk-free mark-to-market and the positive exposure:

$$e^{-}(t) = \mathrm{MtM}(t) - e^{+}(t) = \max(-\mathrm{MtM}(t), 0)$$

  The negative exposure is then the equivalent of the positive exposure from the perspective of the counterparty

The CVA formula

CVA is the risk-neutral discounted expected value of the potential loss:

$$\mathrm{CVA} = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}\left\{\boldsymbol{\tau}_B \leq T\right\} \cdot e^{-\int_0^{\boldsymbol{\tau}_B} r_t \, \mathrm{d}t} \cdot L\right]$$

where:

- T is the maturity of the OTC derivative
- $\boldsymbol{\tau}_B$ is the default time of Bank B
- L is the counterparty loss:

$$L = \left(1 - \boldsymbol{\mathcal{R}}_B\right) \cdot e^{+}\left(\boldsymbol{\tau}_B\right)$$

The CVA formula

Using usual assumptions, we obtain:

$$\text{CVA} = (1 - \mathcal{R}_B) \cdot \int_0^T B_0(t)\,\text{EpE}(t)\,\mathrm{d}\mathbf{F}_B(t)$$

where:

- $\text{EpE}(t)$ is the risk-neutral expected positive exposure:

$$\text{EpE}(t) = \mathbb{E}^{\mathbb{Q}}\left[e^+(t)\right]$$

- $\mathbf{F}_B$ is the cumulative distribution function of $\boldsymbol{\tau}_B$

The CVA formula

Since $\mathbf{S}_B(t) = 1 - \mathbf{F}_B(t)$, we obtain:

$$\text{CVA} = (1 - \mathcal{R}_B) \cdot \int_0^T -B_0(t)\,\text{EpE}(t)\,\mathrm{d}\mathbf{S}_B(t)$$

The DVA formula

The debit value adjustment is defined as the risk-neutral discounted expected value of the potential gain:

$$\mathrm{DVA} = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}\left\{\boldsymbol{\tau}_A \leq T\right\} \cdot e^{-\int_0^{\boldsymbol{\tau}_A} r_t \, \mathrm{d}t} \cdot G\right]$$

where:

- $\boldsymbol{\tau}_A$ is the default time of Bank A
- G is the counterparty gain:

$$G = \left(1 - \boldsymbol{\mathcal{R}}_A\right) \cdot e^-\left(\boldsymbol{\tau}_A\right)$$

The DVA formula

$$\mathrm{DVA} = \left(1 - \boldsymbol{\mathcal{R}}_A\right) \cdot \int_0^T -B_0\left(t\right) \mathrm{EnE}\left(t\right) \, \mathrm{d}\mathbf{S}_A\left(t\right)$$

where $\mathrm{EnE}\left(t\right)$ is the risk-neutral expected negative exposure:

$$\mathrm{EnE}\left(t\right) = \mathbb{E}^{\mathbb{Q}}\left[e^-\left(t\right)\right]$$

The two BCVA formulas

Independent case ($\boldsymbol{\tau}_B \perp \boldsymbol{\tau}_A$)

$$\begin{aligned}
\mathrm{BCVA} &= \mathrm{DVA} - \mathrm{CVA} \\
&= (1 - \boldsymbol{\mathcal{R}}_A) \cdot \int_0^T -B_0(t)\,\mathrm{EnE}(t)\;\mathrm{d}\mathbf{S}_A(t) - \\
&\quad (1 - \boldsymbol{\mathcal{R}}_B) \cdot \int_0^T -B_0(t)\,\mathrm{EpE}(t)\;\mathrm{d}\mathbf{S}_B(t)
\end{aligned}$$

General case

We must consider the joint survival function of $(\boldsymbol{\tau}_A, \boldsymbol{\tau}_B)$:

$$\mathrm{BCVA} = \mathbb{E}^{\mathbb{Q}}\left[\begin{array}{l} \mathbb{1}\{\boldsymbol{\tau}_A \leq \min(T, \boldsymbol{\tau}_B)\} \cdot e^{-\int_0^{\boldsymbol{\tau}_A} r_t\,\mathrm{d}t} \cdot G - \\ \mathbb{1}\{\boldsymbol{\tau}_B \leq \min(T, \boldsymbol{\tau}_A)\} \cdot e^{-\int_0^{\boldsymbol{\tau}_B} r_t\,\mathrm{d}t} \cdot L \end{array}\right]$$

Interpretation of the CVA measure

If we assume that the yield curve is flat and $\mathbf{S}_B(t) = e^{-\lambda_B t}$, we have:

$$\mathrm{d}\mathbf{S}_B(t) = -\lambda_B e^{-\lambda_B t}\,\mathrm{d}t$$

and:

$$\begin{aligned} \mathrm{CVA} &= (1 - \mathcal{R}_B) \cdot \int_0^T e^{-rt}\,\mathrm{EpE}(t)\,\lambda_B e^{-\lambda_B t}\,\mathrm{d}t \\ &= \mathcal{s}_B \cdot \int_0^T e^{-(r+\lambda_B)t}\,\mathrm{EpE}(t)\,\mathrm{d}t \end{aligned}$$

$\Rightarrow$ CVA is the product of the CDS spread and the discounted value of the expected positive exposure

Exercise III

Exercise (HFRM, Exercise 4.4.5, page 303)

We assume that the mark-to-market value is given by:

$$\mathrm{MtM}(t) = N \int_t^T f(t, T) B_t(s) \, \mathrm{d}s - N \int_t^T f(0, T) B_t(s) \, \mathrm{d}s$$

where N and T are the notional and the maturity of the swap, and $f(t, T)$ is the instantaneous forward rate which follows a geometric Brownian motion:

$$\mathrm{d}f(t, T) = \mu f(t, T) \, \mathrm{d}t + \sigma f(t, T) \, \mathrm{d}W(t)$$

We also assume that the yield curve is flat – $B_t(s) = e^{-r(s-t)}$ – and the risk-neutral survival function is $\mathbf{S}(t) = e^{-\lambda t}$

Solution

Solution (Syrkin and Shirazi, 2015; HFRM-CB, Section 4.4.5, pages 82-85)

We have:

$$\mathrm{CVA}(t) = s_B \cdot \int_t^T e^{-(r+\lambda)(u-t)} \, \mathrm{EpE}(u) \, \mathrm{d}u$$

where:

$$\mathrm{EpE}(t) = Nf(0,T)\,\varphi(t,T)\left(e^{\mu t}\Phi\left(\left(\frac{\mu}{\sigma}+\frac{1}{2}\sigma\right)\sqrt{t}\right) - \Phi\left(\left(\frac{\mu}{\sigma}-\frac{1}{2}\sigma\right)\sqrt{t}\right)\right)$$

and:

$$\varphi(t,T) = \frac{1-e^{-r(T-t)}}{r}$$

Numerical example: $N = 1000$, $f(0,T) = 5\%$, $\mu = 2\%$, $\sigma = 25\%$, $T = 10$ years and $\mathcal{R}_B = 50\%$

Solution

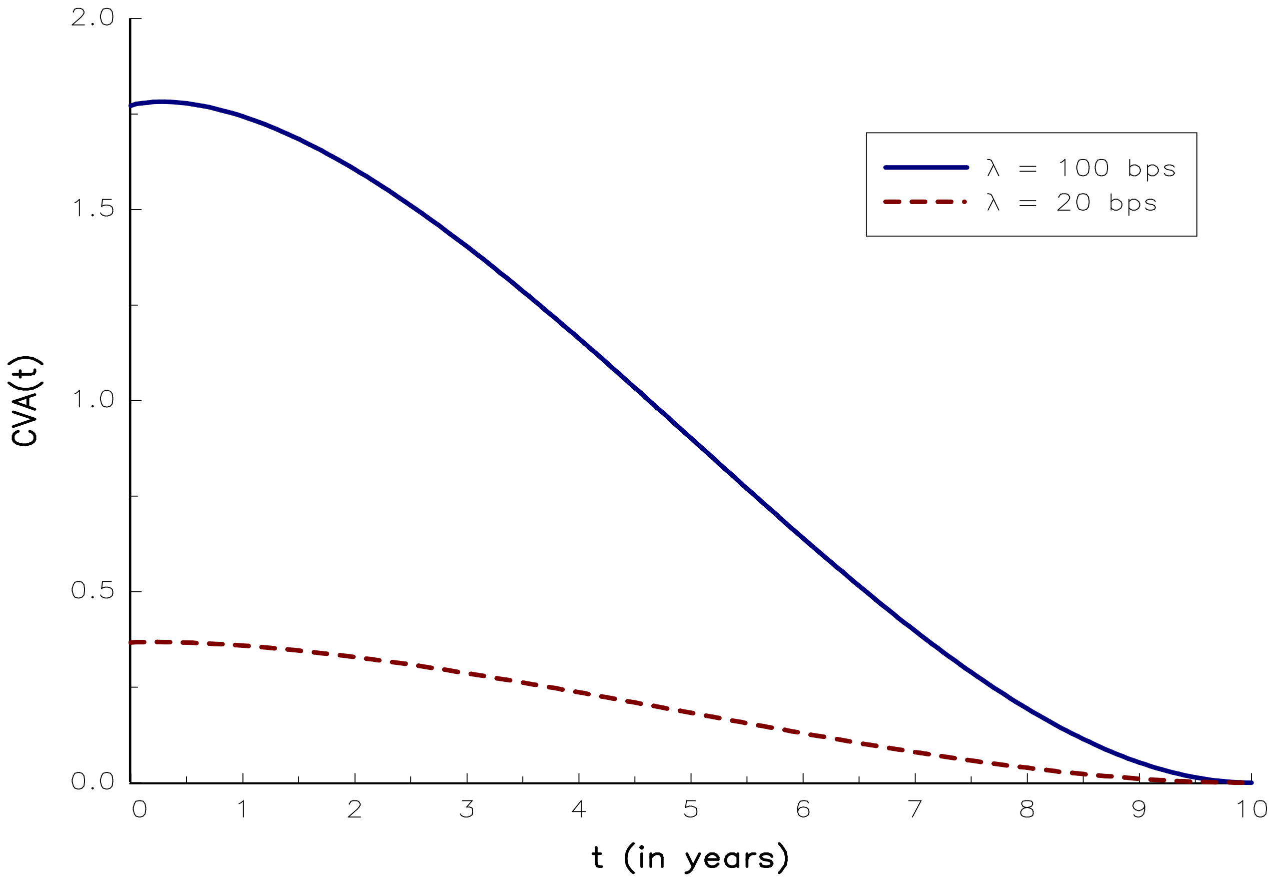


Figure: CVA of fixed-float swaps

Practical implementation

We approximate the integral by a sum:

$$\mathrm{CVA} = (1 - \mathcal{R}_B) \cdot \sum_{t_i \leq T} B_0(t_i) \cdot \mathrm{EpE}(t_i) \cdot (\mathbf{S}_B(t_{i-1}) - \mathbf{S}_B(t_i))$$

and:

$$\mathrm{DVA} = (1 - \mathcal{R}_A) \cdot \sum_{t_i \leq T} B_0(t_i) \cdot \mathrm{EnE}(t_i) \cdot (\mathbf{S}_A(t_{i-1}) - \mathbf{S}_A(t_i))$$

where $\{t_i\}$ is a partition of $[0, T]$

Practical implementation

We have:

$$\mathbf{S}_B(t_{i-1}) - \mathbf{S}_B(t_i) = \Pr\{t_{i-1} < \boldsymbol{\tau}_B \leq t_i\} = \mathrm{PD}_B(t_{i-1}, t_i)$$

$\mathrm{PD}_B(t_{i-1}, t_i)$ is a risk-neutral probability

The credit triangle relationship is:

$$s_B(t) = (1 - \mathcal{R}_B) \cdot \lambda_B(t)$$

We deduce that:

$$\mathbf{S}_B(t) = \exp(-\lambda_B(t) \cdot t) = \exp\left(-\frac{s_B(t) \cdot t}{1 - \mathcal{R}_B}\right)$$

and:

$$\mathrm{PD}_B(t_{i-1}, t_i) = \exp\left(-\frac{s_B(t_{i-1}) \cdot t_{i-1}}{1 - \mathcal{R}_B}\right) - \exp\left(-\frac{s_B(t_i) \cdot t_i}{1 - \mathcal{R}_B}\right)$$

Comparison with AM-CVA (2010 version of Basel III)

BCBS approximates the integral by the middle Riemann sum:

$$\mathrm{CVA} = \mathrm{LGD}_B \cdot \sum_{t_i \leq T} \left(\frac{\mathrm{EpE}\left(t_{i-1}\right) B_0\left(t_{i-1}\right) + B_0\left(t_i\right) \mathrm{EpE}\left(t_i\right)}{2} \right) \cdot \mathrm{PD}_B\left(t_{i-1}, t_i\right)$$

where:

- $\mathrm{LGD} = 1 - \mathcal{R}_B$ is the risk-neutral loss given default of the counterparty B
- $\mathrm{PD}_B\left(t_{i-1}, t_i\right)$ is the risk neutral probability of default between t_{i-1} and t_i:

$$\mathrm{PD}_B\left(t_{i-1}, t_i\right) = \max\left(\exp\left(-\frac{s\left(t_{i-1}\right)}{\mathrm{LGD}_B} \cdot t_{i-1}\right) - \exp\left(-\frac{s\left(t_i\right)}{\mathrm{LGD}_B} \cdot t_i\right), 0\right)$$

Basel III

2010 version of Basel III

- Standardized method (SM-CVA)
- Advanced method (AM-CVA)

2017 version of Basel III

- Basic approach (BA-CVA)
- Standardized approach (SA-CVA)

$\Rightarrow$ The Basel Committee completely flip-flopped within the same accord, since the 2017 version will replace the 2010 version in January 2022

Basic approach (BA-CVA)

The capital requirement is equal to:

$$\mathcal{K} = \beta \cdot \mathcal{K}^{\text{Reduced}} + (1 - \beta) \cdot \mathcal{K}^{\text{Hedged}}$$

where $\mathcal{K}^{\text{Reduced}}$ and $\mathcal{K}^{\text{Hedged}}$ are the capital requirements without and with hedging recognition

- The reduced version of the BA-CVA is obtained by setting β to 100%
- A bank that actively hedges CVA risks may choose the full version of the BA-CVA and $\beta = 25\%$

Reduced version

We have:

$$\mathcal{K}^{\mathrm{Reduced}} = \sqrt{\left(\rho \cdot \sum_j \mathrm{SCVA}_j\right)^2 + (1 - \rho^2) \cdot \sum_j \mathrm{SCVA}_j^2}$$

where:

- $\rho = 50\%$
- SCVA_j is the CVA capital requirement for the j^{th} counterparty:

$$\mathrm{SCVA}_j = \frac{1}{\alpha} \cdot \mathrm{RW}_j \cdot \sum_k \mathrm{DF}_k \cdot \mathrm{EAD}_k \cdot \mathrm{M}_k$$

- $\alpha = 1.4$
- RW_j is the risk weight for counterparty j
- k is the netting set, DF_k is the discount factor, EAD_k is the CCR exposure at default, M_k is the effective maturity

Reduced version

RW_j depends on the credit quality of the counterparty (IG/HY) and its sector:

Table: Supervisory risk weights (BA-CVA)

Sector	Credit quality IG	Credit quality HY/NR
Sovereign	0.5%	3.0%
Local government	1.0%	4.0%
Financial	5.0%	12.0%
Basic material, energy, industrial, agriculture, manufacturing, mining and quarrying	3.0%	7.0%
Consumer goods and services, transportation and storage, administrative and support service activities	3.0%	8.5%
Technology, telecommunication	2.0%	5.5%
Health care, utilities, professional and technical activities	1.5%	5.0%
Other sector	5.0%	12.0%

Hedged version

The full version of the BA-CVA recognizes hedging instruments (single-name CDS and index CDS):

$$\mathcal{K}^{\mathrm{Hedged}} = \sqrt{K_1 + K_2 + K_3}$$

where:

1. K_1 aggregates the systematic risk components of the CVA risk:
$$K_1 = \left(\rho \cdot \sum_j (\mathrm{SCVA}_j - \mathrm{SNH}_j) - \mathrm{IH}\right)^2$$
2. K_2 aggregates the idiosyncratic risk components of the CVA risk:
$$K_2 = \left(1 - \rho^2\right) \cdot \sum_j (\mathrm{SCVA}_j - \mathrm{SNH}_j)^2$$
3. K_3 corresponds to the hedging misalignment risk because of the mismatch between indirect and single-name hedges:
$$K_3 = \sum_j \mathrm{HMA}_j$$

Hedged version

Single-name hedging

SNH_j is the CVA reduction for counterparty j due to single-name hedging

$$\mathrm{SNH}_j = \sum_{h \in j} \varrho_{h,j} \cdot \left(\mathrm{RW}_h \cdot \mathrm{DF}_h \cdot N_h \cdot \mathrm{M}_h\right)$$

where:

- h represents the single-name CDS transaction, $\varrho_{h,j}$ is the supervisory correlation, DF_h is the discount factor, N_h is the notional and M_h is the remaining maturity
- These quantities are calculated at the single-name CDS level
- The correlation $\varrho_{h,j}$ between the credit spread of the counterparty and the credit spread of the CDS can take three values:
  1. 100% if CDS h directly refers to counterparty j
  2. 80% if CDS h has a legal relation with counterparty j
  3. 50% if CDS h and counterparty j are of the same sector and region

Hedged version
Index hedging

IH is the global CVA reduction due to index hedging:

$$\mathrm{IH} = \sum_{h'} \mathrm{RW}_{h'} \cdot \mathrm{DF}_{h'} \cdot N_{h'} \cdot \mathrm{M}_{h'}$$

where:

- h' represents the index CDS transaction
- The risk weight is the weighted average of risk weights of RW_j:

$$\mathrm{RW}_{h'} = 0.7 \cdot \sum_{j \in h'} w_j \cdot \mathrm{RW}_j$$

  where w_j is the weight of the counterparty/sector j in the index CDS h'

Hedged version

Hedging mismatch

The hedging misalignment risk is equal to:

$$\mathrm{HMA}_j = \sum_{h \in j} \left(1 - \varrho_{h,j}^2\right) \cdot \left(\mathrm{RW}_h \cdot \mathrm{DF}_h \cdot N_h \cdot \mathrm{M}_h\right)^2$$

Basic approach (BA-CVA)
Special cases

If there is no hedge, we have $\mathrm{SNH}_j = 0$, $\mathrm{HMA}_j = 0$, $\mathrm{IH} = 0$, and

$$\mathcal{K} = \mathcal{K}^{\mathrm{Reduced}}$$

If there is no hedging misalignment risk and no index CDS hedging, we have:

$$\mathcal{K} = \sqrt{\left(\rho \cdot \sum_j \mathcal{K}_j\right)^2 + (1 - \rho^2) \cdot \sum_j \mathcal{K}_j^2}$$

where $\mathcal{K}_j = \mathrm{SCVA}_j - \mathrm{SNH}_j$ is the single-name capital requirement for counterparty j

Exercise IV

Exercise

We assume that the bank has three financial counterparties A, B and C, that are respectively rated IG, IG and HY. There are 4 OTC transactions, whose characteristics are the following:

Transaction k	1	2	3	4
Counterparty	A	A	B	C
EAD_k	100	50	70	20
M_k	1	1	0.5	0.5

In order to reduce the counterparty credit risk, the bank has purchased a CDS protection on A for an amount of \$75 mn, a CDS protection on B for an amount of \$10 mn and a HY Financial CDX for an amount of \$10 mn. The maturity of hedges exactly matches the maturity of transactions. However, the CDS protection on B is indirect, because the underlying name is not B, but B' which is the parent company of B

Solution ($\mathcal{K}^{\mathrm{Reduced}}$)

- We calculate the discount factors DF_k for the four transactions: $\mathrm{DF}_1 = \mathrm{DF}_2 = 0.9754$ and $\mathrm{DF}_3 = \mathrm{DF}_4 = 0.9876$
- We calculate the single-name capital for each counterparty:

$$\begin{aligned}\mathrm{SCVA}_A &= \frac{1}{\alpha} \times \mathrm{RW}_A \times (\mathrm{DF}_1 \times \mathrm{EAD}_1 \times \mathrm{M}_1 + \mathrm{DF}_2 \times \mathrm{EAD}_2 \times \mathrm{M}_2) \\ &= \frac{1}{1.4} \times 5\% \times (0.9754 \times 100 \times 1 + 0.9754 \times 50 \times 1) \\ &= 5.225\end{aligned}$$

  We also find that $\mathrm{SCVA}_B = 1.235$ and $\mathrm{SCVA}_C = 0.847$
- It follows that $\sum_j \mathrm{SCVA}_j = 7.306$ and $\sum_j \mathrm{SCVA}_j^2 = 29.546$
- The capital requirement without hedging is equal to:

$$\mathcal{K}^{\mathrm{Reduced}} = \sqrt{(0.5 \times 7.306)^2 + (1 - 0.5^2) \times 29.546} = 5.959$$

Solution ($\mathcal{K}^{\mathrm{Hedged}}$)

- We calculate the single-name hedge parameters:

$$\mathrm{SNH}_A = 5\% \times 100\% \times 0.9754 \times 75 \times 1 = 3.658$$

and:

$$\mathrm{SNH}_B = 5\% \times 80\% \times 0.9876 \times 10 \times 0.5 = 0.198$$

- Since the CDS protection is on B' and not B, there is a hedging misalignment risk:

$$\mathrm{HMA}_B = 0.05^2 \times \left(1 - 0.80^2\right) \times (0.9876 \times 10 \times 0.5)^2 = 0.022$$

- For the CDX protection, we have:

$$\mathrm{IH} = (0.7 \times 12\%) \times 0.9876 \times 10 \times 0.5 = 0.415$$

- We obtain $K_1 = 1.718$, $K_2 = 3.187$, $K_3 = 0.022$ and

$$\mathcal{K}^{\mathrm{Hedged}} = \sqrt{1.718^2 + 3.187^2 + 0.022^2} = 2.220$$

Solution (regulatory capital)

The capital requirement is equal to \$3.154 mn:

$$\mathcal{K} = 0.25 \times 5.959 + 0.75 \times 2.220 = 3.154$$

Standardized approach (SA-CVA)

Remark

SA-CVA ≈ SA-TB

$$\mathcal{K} = \mathcal{K}^{\mathrm{Delta}} + \mathcal{K}^{\mathrm{Vega}}$$

- Two portfolios:
  1. The CVA portfolio
  2. The hedging portfolio
- For each risk (delta and vega), we calculate the weighted CVA sensitivity of each risk factor $\mathcal{F}_j$:

$$\mathrm{WS}_j^{\mathrm{CVA}} = S_j^{\mathrm{CVA}} \cdot \mathrm{RW}_j$$

and:

$$\mathrm{WS}_j^{\mathrm{Hedge}} = S_j^{\mathrm{Hedge}} \cdot \mathrm{RW}_j$$

where S_j and RW_j are the net sensitivity of the CVA or hedging portfolio with respect to the risk factor and the risk weight of $\mathcal{F}_j$

Standardized approach (SA-CVA)

- We aggregate the weighted sensitivity in order to obtain a net figure:

$$\mathrm{WS}_j = \mathrm{WS}_j^{\mathrm{CVA}} + \mathrm{WS}_j^{\mathrm{Hedge}}$$

- We calculate the capital requirement for the risk bucket $\mathcal{B}_k$:

$$\mathcal{K}_{\mathcal{B}_k} = \sqrt{\sum_j \mathrm{WS}_j^2 + \sum_{j' \neq j} \rho_{j,j'} \cdot \mathrm{WS}_j \cdot \mathrm{WS}_{j'} + 1\% \cdot \sum_j \left(\mathrm{WS}_j^{\mathrm{Hedge}}\right)^2}$$

where $\mathcal{F}_j \in \mathcal{B}_k$

- We aggregate the different buckets for a given risk class:

$$\mathcal{K}^{\mathrm{Delta/Vega}} = m_{\mathrm{CVA}} \cdot \sqrt{\sum_k \mathcal{K}_{\mathcal{B}_k}^2 + \sum_{k' \neq k} \gamma_{k,k'} \cdot \mathcal{K}_{\mathcal{B}_k} \cdot \mathcal{K}_{\mathcal{B}_{k'}}}$$

where $m_{\mathrm{CVA}} = 1.25$ is the multiplier factor

CVA and wrong/right way risk

- CVA trading desk
- How to be sure that the CVA hedging portfolio does not create itself another source of hidden wrong way risk?
- In practice, **market and credit risks are correlated!**
- Two approaches
  1. The copula model (Cespedes *et al.*, 2010)
  2. The hazard rate model (Hull and White, 2012)

Exposure at default

- In the case of a margin agreement, the counterparty needs to post collateral and the exposure at default becomes:

$$e^{+}(t) = \max\left(\mathrm{MtM}(t) - C(t), 0\right)$$

  where $C(t)$ is the collateral value at time t
- The collateral transfer occurs when the mark-to-market exceeds a threshold H:

$$C(t) = \max\left(\mathrm{MtM}(t - \delta_C) - H, 0\right)$$

  where:
  - H is the minimum collateral transfer amount
  - $\delta_C \geq 0$ is the margin period of risk (MPOR)
- We obtain:

$$\begin{aligned} e^{+}(t) &= \mathrm{MtM}(t) \cdot \mathbb{1}\left\{0 \leq \mathrm{MtM}(t), \mathrm{MtM}(t - \delta_C) < H\right\} + \\ & \quad \left(\mathrm{MtM}(t) - \mathrm{MtM}(t - \delta_C) + H\right) \cdot \\ & \quad \mathbb{1}\left\{H \leq \mathrm{MtM}(t - \delta_C) \leq \mathrm{MtM}(t) + H\right\} \end{aligned}$$

Special cases

- When $H = +\infty$, $C(t)$ is equal to zero and we obtain:

$$e^{+}(t) = \max(\mathrm{MtM}(t), 0)$$

- When $H = 0$, the collateral $C(t)$ is equal to $\mathrm{MtM}(t - \delta_C)$ and the counterparty exposure becomes:

$$e^{+}(t) = \max(\mathrm{MtM}(t) - \mathrm{MtM}(t - \delta_C), 0) = \max(\mathrm{MtM}(t - \delta_C, t), 0)$$

  The CCR corresponds to the variation of the mark-to-market $\mathrm{MtM}(t - \delta_C, t)$ during the liquidation period $[t - \delta_C, t]$

- When δ_C is set to zero, we deduce that:

$$e^{+}(t) = \mathrm{MtM}(t) \cdot \mathbb{1}\{0 \leq \mathrm{MtM}(t) < H\} + H \cdot \mathbb{1}\{H \leq \mathrm{MtM}(t)\}$$

- When δ_C is set to zero and there is no minimum collateral transfer amount, the counterparty credit risk vanishes:

$$e^{+}(t) = 0$$

Illustration

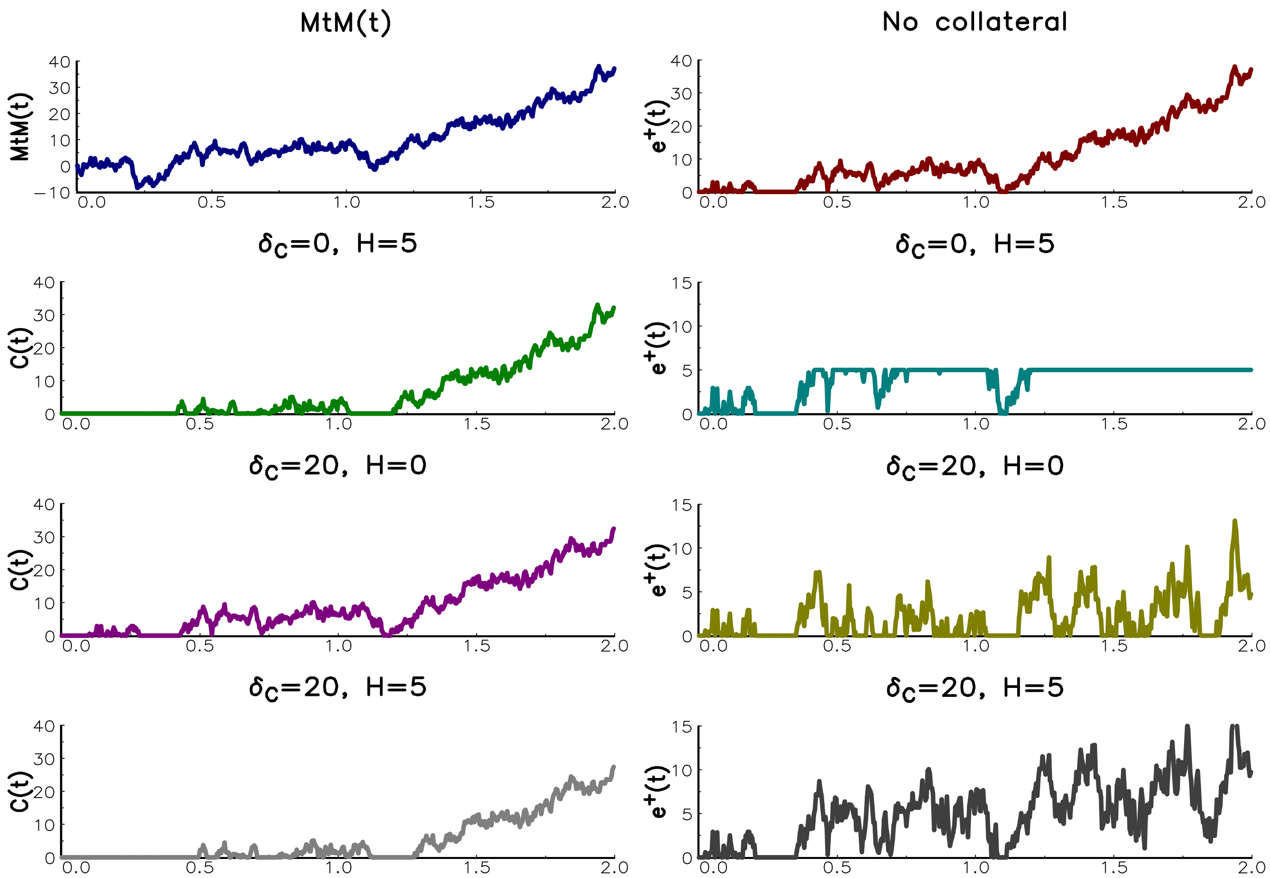


Figure: Impact of collateral on the counterparty exposure

Collateral risk management

Two ways to reduce the counterparty risk:

1. Reducing the haircut ($H \searrow 0$)
2. Reducing the margin period of risk ($\delta_C \searrow 0$)

Trade-off between risk and operational cost & process

Risk allocation

We recall the Euler allocation principle:

$$\mathcal{R}(w) = \sum_{i=1}^{n} \mathcal{RC}_i = \sum_{i=1}^{n} w_i \cdot \frac{\partial\, \mathcal{R}(w)}{\partial\, w_i}$$

Risk allocation

Application to a CVA portfolio

$$\mathrm{CVA}(w) = (1 - \mathcal{R}_B) \cdot \int_0^T -B_0(t)\,\mathrm{EpE}(t; w)\,\mathrm{d}\mathbf{S}_B(t)$$

where $\mathrm{EpE}(t; w)$ is the expected positive exposure with respect to the portfolio w. The Euler allocation principle becomes:

$$\mathrm{CVA}(w) = \sum_{i=1}^{n} \mathrm{CVA}_i(w)$$

where $\mathrm{CVA}_i(w)$ is the CVA risk contribution of the i^{th} component:

$$\mathrm{CVA}_i(w) = (1 - \mathcal{R}_B) \cdot \int_0^T -B_0(t)\,\mathrm{EpE}_i(t; w)\,\mathrm{d}\mathbf{S}_B(t)$$

and $\mathrm{EpE}_i(t; w)$ is the EpE risk contribution of the i^{th} component

Risk allocation

What is the challenge?

Computing the EpE risk contribution:

$$\mathrm{EpE}_i\left(t;w\right)=w_i\cdot\frac{\partial\ \mathrm{EpE}\left(t;w\right)}{\partial\ w_i}$$

Very difficult and almost impossible ⇒ needs simplification

Exercises

- Counterparty credit risk (CCR)
  - Exercise 4.4.1 – Impact of netting agreements in counterparty credit risk
  - Exercise 4.4.2 – Calculation of the effective expected positive exposure
  - Exercise 4.4.3 – Calculation of the capital charge for counterparty credit risk
- Credit valuation adjustment (CVA)
  - Exercise 4.4.4 – Calculation of CVA and DVA measures
  - Exercise 4.4.5 – Approximation of the CVA for an interest rate swap

References

Basel Committee on Banking Supervision (2006)
International Convergence of Capital Measurement and Capital Standards — A Revised Framework — Comprehensive version, June 2006.

Basel Committee on Banking Supervision (2014)
The Standardized Approach for Measuring Counterparty Credit Risk Exposures, March 2014 (revision April 2014).

Basel Committee on Banking Supervision (2017)
Basel III: Finalising Post-crisis Reforms, December 2017.

References

Canabarro, E., and Duffie, D. (2003)
Measuring and Marking Counterparty Risk, Chapter 9 in Tilman, L. (Ed.), *Asset/Liability Management for Financial Institutions*, Institutional Investor Books
http://web.stanford.edu/~duffie/Chapter_09.pdf.

European Banking Authority (2015)
EBA Report on CVA, February 2015.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 4.

Roncalli, T. (2020)
Handbook of Financial Risk Management – Companion Book, Chapter 4.

Course 2023-2024 in Financial Risk Management
Lecture 5. Operational Risk

Thierry Roncalli*

*Amundi Asset Management[14]

*University of Paris-Saclay

September 2023

[14]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- **Lecture 5: Operational Risk**
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

A long list of operational risk losses:

- 1983: Banco Ambrosiano Vatican Bank (money laundering)
- 1995: Barings (rogue trading)
- 1996: Summitomo Bank (rogue trading)
- 1996: Crédit Lyonnais (headquarter fire)
- Etc.

Since the end of the nineties, new themes: operational risk, legal risk, compliance, money laundering, etc.

Definition

Definition

The Basel Committee defines the operational risk in the following way:

> *"Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk"*

Loss event type classification

1. Internal fraud ("*losses due to acts of a type intended to defraud, misappropriate property or circumvent regulations, the law or company policy, excluding diversity/discrimination events, which involves at least one internal party*")
   1. Unauthorized activity
   2. Theft and fraud
2. External fraud ("*losses due to acts of a type intended to defraud, misappropriate property or circumvent the law, by a third party*")
   1. Theft and fraud
   2. Systems security
3. Employment practices and workplace safety ("*losses arising from acts inconsistent with employment, health or safety laws or agreements, from payment of personal injury claims, or from diversity/discrimination events*")
   1. Employee relations
   2. Safe environment
   3. Diversity & discrimination

Loss event type classification

4. Clients, products & business practices ("*losses arising from an unintentional or negligent failure to meet a professional obligation to specific clients (including fiduciary and suitability requirements), or from the nature or design of a product*")
    1. Suitability, disclosure & fiduciary
    2. Improper business or market practices
    3. Product flaws
    4. Selection, sponsorship & exposure
    5. Advisory activities
5. Damage to physical assets ("*losses arising from loss or damage to physical assets from natural disaster or other events*")
    1. Disasters and other events

Loss event type classification

6. Business disruption and system failures (“*losses arising from disruption of business or system failures*”)
    1. Systems
7. Execution, delivery & process management (“*losses from failed transaction processing or process management, from relations with trade counterparties and vendors*”)
    1. Transaction capture, execution & maintenance
    2. Monitoring and reporting
    3. Customer intake and documentation
    4. Customer/client account management
    5. Trade counterparties
    6. Vendors & suppliers

Loss data collection exercise (LDCE)

Table: Internal losses larger than €20 000 per year

Year	pre 2002	2002	2003	2004	2005	2006	2007
n_L	14 017	10 216	13 691	22 152	33 216	36 386	36 622
L (in € bn)	3.8	12.1	4.6	7.2	9.7	7.4	7.9
n_B	24	35	55	68	108	115	117

More and more operational risk losses:

- Société Générale in 2008 ($7.2 bn), Morgan Stanley in 2008 ($9.0 bn), BPCE in 2008 ($1.1 bn), UBS in 2011 ($2 bn), JPMorgan Chase in 2012 ($5.8 bn), etc.
- Libor scandal: $2.5 bn for Deutsche Bank, $1 bn for Rabobank, $545 mn for UBS, etc.
- Forex scandal: six banks (BoA, Barclays, Citi, JPM, UBS and RBS) agreed to pay fines totaling $5.6 bn in May 2015
- BNP Paribas payed a fine of $8.9 bn in June 2014 (anti-money laundering control)
- Etc.

Basel II versus Basel III

Basel II

- Basic indicator approach (BIA)
- The standardized approach (TSA)
- Advanced measurement approaches (AMA)

Basel III

- Standardized approach (SA-OR)
- Pillar II

Basic indicator approach (BIA)

The capital charge is a fixed percentage of annual gross income:

$$\mathcal{K} = \alpha \cdot \overline{\mathrm{GI}}$$

where $\alpha = 15\%$ and $\overline{\mathrm{GI}}$ is the average of the positive gross income over the previous three years:

$$\overline{\mathrm{GI}} = \frac{\max\left(\mathrm{GI}_{t-1}, 0\right) + \max\left(\mathrm{GI}_{t-2}, 0\right) + \max\left(\mathrm{GI}_{t-3}, 0\right)}{\sum_{k=1}^{3} \mathbb{1}\left\{\mathrm{GI}_{t-k} > 0\right\}}$$

The standardized approach (TSA)

TSA is an extended version of BIA:

$$\mathcal{K}_{j,t} = \beta_j \cdot \mathrm{GI}_{j,t}$$

where β_j and $\mathrm{GI}_{j,t}$ are a fixed percentage and the gross income corresponding to the j^{th} business line. The total capital charge is the three-year average of the sum of all the capital charges:

$$\mathcal{K} = \frac{1}{3} \sum_{k=1}^{3} \max \left(\sum_{j=1}^{8} \mathcal{K}_{j,t-k}, 0 \right)$$

If the values of gross income are all positive, the total capital charge becomes:

$$\mathcal{K} = \frac{1}{3} \sum_{k=1}^{3} \sum_{j=1}^{8} \beta_j \cdot \mathrm{GI}_{j,t-k} = \sum_{j=1}^{8} \beta_j \cdot \overline{\mathrm{GI}}_j$$

The standardized approach (TSA)

Table: Mapping of business lines for operational risk

Level 1	Level 2	β_j
Corporate Finance	Corporate Finance Municipal/Government Finance Merchant Banking Advisory Services	18%
Trading & Sales	Sales Market Making Proprietary Positions Treasury	18%
Retail Banking	Retail Banking Private Banking Card Services	12%
Commercial Banking[♮]	Commercial Banking	12%
Payment & Settlement	External Clients	18%
Agency Services	Custody Corporate Agency Corporate Trust	15%
Asset Management	Discretionary Fund Management Non-Discretionary Fund Management	12%
Retail Brokerage	Retail Brokerage	12%

The standardized approach (TSA)

What is the difference between corporate finance, trading & sales and commercial banking?

- **Corporate finance**: mergers and acquisitions, underwriting, securitization, syndications, IPO, debt placements
- **Trading & sales**: buying and selling of securities and derivatives, own position securities, lending and repos, brokerage
- **Commercial banking**: project finance, real estate, export finance, trade finance, factoring, leasing, lending, guarantees, bills of exchange

Advanced measurement approaches (AMA)

The AMA method is defined by certain criteria without refereing to a specific statistical model:

- The capital charge should cover the one-year operational loss at the 99.9% confidence level $(\mathrm{UL} + \mathrm{EL})$
- A minimum five-year observation period of internal loss data
- The model can incorporate the risk mitigation impact of insurance, which is limited to 20% of the total operational risk capital charge

Advanced measurement approaches (AMA)

Table: Distribution of annualized operational losses (in %)

Business line	Event type 1	2	3	4	5	6	7	All
Corporate Finance	0.2	0.1	0.6	93.7	0.0	0.0	5.4	28.0
Trading & Sales	11.0	0.3	2.3	29.0	0.2	1.8	55.3	13.6
Retail Banking	6.3	19.4	9.8	40.4	1.1	1.5	21.4	32.0
Commercial Banking	11.4	15.2	3.1	35.5	0.4	1.7	32.6	7.6
Payment & Settlement	2.8	7.1	0.9	7.3	3.2	2.3	76.4	2.6
Agency Services	1.0	3.2	0.7	36.0	18.2	6.0	35.0	2.6
Asset Management	11.1	1.0	2.5	30.8	0.3	1.5	52.8	2.5
Retail Brokerage	18.1	1.4	6.3	59.5	0.1	0.2	14.4	5.1
Unallocated	6.5	2.8	28.4	28.3	6.5	1.3	26.2	6.0
All	6.1	8.0	6.0	52.4	1.4	1.2	24.9	100.0

Basel III (SA-OR or SMA)

Remark

The standardized measurement approach (SMA) will replace the three approaches of the Basel II framework in 2022. AMA may be used for Pillar 2

The SMA is based on three components:

1. Business indicator (BI)
2. Business indicator component (BIC)
3. Internal loss multiplier (ILM)

Basel III (SA-OR or SMA)

- The business indicator is a proxy of the operational risk:

$$\mathrm{BI} = \mathrm{ILDC} + \mathrm{SC} + \mathrm{FC}$$

where ILDC is the interest, leases and dividends component, SC is the services component and FC is the financial component. The underlying idea is to list the main activities that generate operational risk:

$$\left\{ \begin{array}{l} \mathrm{ILDC} = \min\left(\left|\mathrm{INC} - \mathrm{EXP}\right|, 2.25\% \cdot \mathrm{IRE}\right) + \mathrm{DIV} \\ \mathrm{SC} = \max\left(\mathrm{OI}, \mathrm{OE}\right) + \max\left(\mathrm{FI}, \mathrm{FE}\right) \\ \mathrm{FC} = \left|\Pi_{\mathrm{TB}}\right| + \left|\Pi_{\mathrm{BB}}\right| \end{array} \right.$$

where INC represents the interest income, EXP the interest expense, IRE the interest earning assets, DIV the dividend income, OI the other operating income, OE the other operating expense, FI the fee income, FE the fee expense, Π_{TB} the net P&L of the trading book and Π_{BB} the net P&L of the banking book

Basel III (SA-OR or SMA)

- The business indicator component is given by:

$$\mathrm{BIC} = 12\% \cdot \min\left(\mathrm{BI}, \$1 \text{ bn}\right) + 15\% \cdot \min\left(\mathrm{BI} - 1, \$30 \text{ bn}\right) + 18\% \cdot \min\left(\mathrm{BI} - 30\right)^{+}$$

- The internal loss multiplier is equal to:

$$\mathrm{ILM} = \ln\left(e^{1} - 1 + \left(\frac{15 \cdot \bar{L}}{\mathrm{BIC}}\right)^{0.8}\right)$$

where $\bar{L}$ is the average annual operational risk losses over the last 10 years

- The capital charge for the operational risk is then equal to:

$$\mathcal{K} = \mathrm{ILM} \cdot \mathrm{BIC}$$

LDA and operational risk

The operational risk loss L of the bank is divided into a matrix of homogenous losses:

$$L = \sum_{k=1}^{K} S_k$$

where S_k is the sum of losses of the k^{th} cell and K is the number of cells in the matrix (Basel II $= 7 \times 8 = 56$ cells)

Definition

Definition

LDA is a method to model the random loss S_k of a particular cell. It assumes that S_k is the random sum of homogeneous individual losses:

$$S_k = \sum_{n=1}^{N_k(t)} X_n^{(k)}$$

where $N_k(t)$ is the random number of individual losses for the period $[0, t]$ and $X_n^{(k)}$ is the n^{th} individual loss

Two sources of uncertainty:

1. We don't know what will be the magnitude of each loss event (severity risk)
2. We don't know how many losses will occur in the next year (frequency risk)

Assumptions

We consider the random sum:

$$S = \sum_{n=1}^{N(t)} X_n$$

The loss distribution approach is based on the following assumptions:

- The number of losses $N(t)$ follows the loss frequency distribution $\mathbf{P}$
- The sequence of individual losses X_n is independent and identically distributed (*iid*)
- The corresponding probability distribution $\mathbf{F}$ is called the loss severity distribution
- The number of events is independent from the amount of loss events

The probability distribution G of S is the compound distribution $(\mathbf{P}, \mathbf{F})$

Exercise I

Exercise

We assume that the number of losses is distributed as follows:

n	0	1	2	3
$p(n)$	50%	30%	17%	3%

The loss amount can take the values \$100 and \$200 with probabilities 70% and 30%

Show that:

s	0	100	200	300	400	500	600
$\Pr\{S = s\}$	50%	21%	17.33%	8.169%	2.853%	0.567%	0.081%

Compound distribution

The cumulative distribution function of S can be written as:

$$\mathbf{G}(s) = \begin{cases} \sum_{n=1}^{\infty} p(n) \mathbf{F}^{n\star}(s) & \text{for} \quad s > 0 \\ p(0) & \text{for} \quad s = 0 \end{cases}$$

where $\mathbf{F}^{n\star}$ is the n-fold convolution of $\mathbf{F}$ with itself:

$$\mathbf{F}^{n\star}(s) = \Pr\left\{\sum_{i=1}^{n} X_i \leq s\right\}$$

Compound distribution

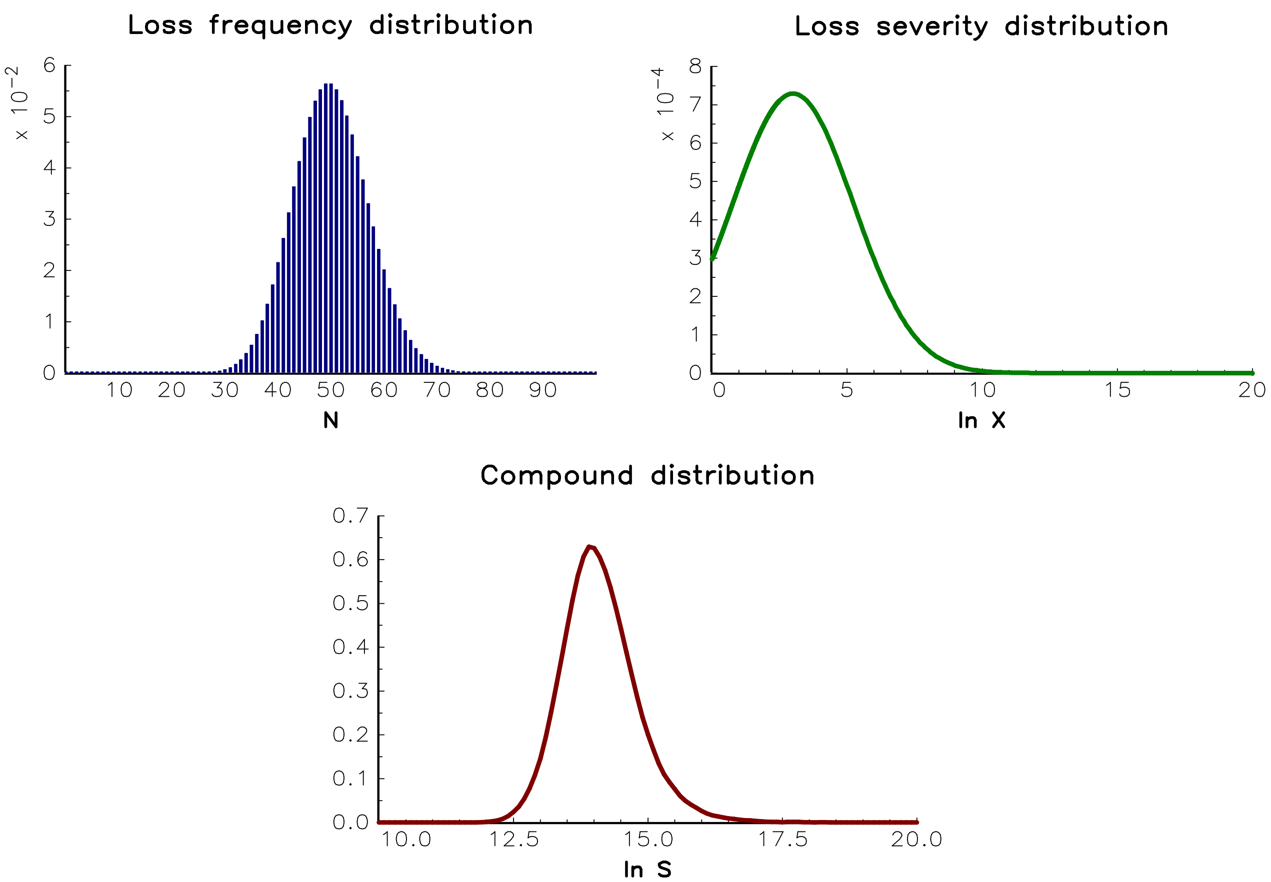


Figure: Compound distribution when $N \sim \mathcal{P}(50)$ and $X \sim \mathcal{LN}(8, 5)$

Regulatory capital

The capital charge (or the capital-at-risk) corresponds to the percentile α:

$$\mathrm{CaR}\left(\alpha\right) = \mathbf{G}^{-1}\left(\alpha\right)$$

The regulatory capital is obtained by setting $\alpha = 99.9\%$:

$$\mathcal{K} = \mathrm{CaR}\left(99.9\%\right)$$

Here are the different steps to implement the loss distribution approach:

- for each cell of the operational risk matrix, we estimate the loss frequency distribution and the loss severity distribution
- we calculate the capital-at-risk
- we define the copula function between the different cells of the operational risk matrix, and deduce the aggregate capital-at-risk

Estimation of the loss severity distribution

Let $\{x_1, \ldots, x_T\}$ the sample collected for a given cell of the operational risk matrix. We consider that the individual losses follow a given parametric distribution $\mathbf{F}$:

$$X \sim \mathbf{F}(x; \theta)$$

where θ is the vector of parameters

The goal is to estimate θ (and $\mathbf{F}$)

Two issues:

- The choice of $\mathbf{F}$
- The choice of the estimation method

Some candidates for the loss severity distribution

- Gamma $X \sim \mathcal{G}(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$
$$\mathbf{F}(x;\theta) = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}$$
- Log-gamma $X \sim \mathcal{LG}(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$
$$\mathbf{F}(x;\theta) = \frac{\gamma(\alpha, \beta \ln x)}{\Gamma(\alpha)}$$
- Log-logistic $X \sim \mathcal{LL}(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$
$$\mathbf{F}(x;\theta) = \frac{1}{1 + (x/\alpha)^{-\beta}} = \frac{x^\beta}{\alpha^\beta + x^\beta}$$
- Log-normal $X \sim \mathcal{LN}(\mu, \sigma^2)$ where $x > 0$ and $\sigma > 0$
$$\mathbf{F}(x;\theta) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$
- Generalized extreme value $X \sim \mathcal{GEV}(\mu, \sigma, \xi)$ where $x > \mu - \sigma/\xi$, $\sigma > 0$ and $\xi > 0$
$$\mathbf{F}(x;\theta) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$
- Pareto $X \sim \mathcal{P}(\alpha, x_-)$ where $x \geq x_-$, $\alpha > 0$ and $x_- > 0$
$$\mathbf{F}(x;\theta) = 1 - \left({}^{x}/_{x_-}\right)^{-\alpha}$$

Some candidates for the loss severity distribution

Table: Density function, mean and variance of parametric probability distribution

Distribution	$f(x;\theta)$	$\mathbb{E}[X]$	$\operatorname{var}(X)$
$\mathcal{G}(\alpha,\beta)$	$\dfrac{\beta^\alpha x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)}$	$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$
$\mathcal{LG}(\alpha,\beta)$	$\dfrac{\beta^\alpha(\ln x)^{\alpha-1}}{x^{\beta+1}\Gamma(\alpha)}$	$\left(\dfrac{\beta}{\beta-1}\right)^\alpha$ if $\beta>1$	$\left(\dfrac{\beta}{\beta-2}\right)^\alpha-\left(\dfrac{\beta}{\beta-1}\right)^{2\alpha}$ if $\beta>2$
$\mathcal{LL}(\alpha,\beta)$	$\dfrac{\beta(x/\alpha)^{\beta-1}}{\alpha\left(1+(x/\alpha)^\beta\right)^2}$	$\dfrac{\alpha\pi}{\beta\sin(\pi/\beta)}$ if $\beta>1$	$\alpha^2\left(\dfrac{2\pi}{\beta\sin(2\pi/\beta)}-\dfrac{\pi^2}{\beta^2\sin^2(\pi/\beta)}\right)$ if $\beta>2$
$\mathcal{LN}(\mu,\sigma^2)$	$\dfrac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\dfrac{1}{2}\left(\dfrac{\ln x-\mu}{\sigma}\right)^2\right)$	$\exp\left(\mu+\dfrac{1}{2}\sigma^2\right)$	$\exp(2\mu+\sigma^2)(\exp(\sigma^2)-1)$
$\mathcal{GEV}(\mu,\sigma,\xi)$	$\dfrac{1}{\sigma}\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{-(1+1/\xi)}$ $\exp\left\{-\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$	$\mu+\dfrac{\sigma}{\xi}(\Gamma(1-\xi)-1)$ if $\xi<1$	$\dfrac{\sigma^2}{\xi^2}\left(\Gamma(1-2\xi)-\Gamma^2(1-\xi)\right)$ if $\xi<\frac{1}{2}$
$\mathcal{P}(\alpha,x_-)$	$\dfrac{\alpha x_-^\alpha}{x^{\alpha+1}}$	$\dfrac{\alpha x_-}{\alpha-1}$ if $\alpha>1$	$\dfrac{\alpha x_-^2}{(\alpha-1)^2(\alpha-2)}$ if $\alpha>2$

Estimation methods

Method of maximum likelihood (HFRM, Section 10.1.2, page 614)

The log-likelihood function associated to the sample is:

$$\ell(\theta) = \sum_{i=1}^{T} \ln f(x_i; \theta)$$

where $f(x; \theta)$ is the density function

Generalized method of moments (HFRM, Section 10.1.3, page 628)

The empirical moments are:

$$\begin{cases} h_{i,1}(\theta) = x_i - \mathbb{E}[X] \\ h_{i,2}(\theta) = (x_i - \mathbb{E}[X])^2 - \operatorname{var}(X) \end{cases}$$

Estimation methods

If we consider that $X \sim \mathcal{LN}\left(\mu, \sigma^2\right)$, the log-likelihood function is:

$$\ell(\theta) = -\sum_{i=1}^{T} \ln x_i - \frac{T}{2} \ln \sigma^2 - \frac{T}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^{T} \left(\frac{\ln x_i - \mu}{\sigma} \right)^2$$

whereas the empirical moments are:

$$\left\{ \begin{array}{l} h_{i,1}(\theta) = x_i - e^{\mu + \frac{1}{2}\sigma^2} \\ h_{i,2}(\theta) = \left(x_i - e^{\mu + \frac{1}{2}\sigma^2} \right)^2 - e^{2\mu + \sigma^2} \left(e^{\sigma^2} - 1 \right) \end{array} \right.$$

Estimation methods

Example

We assume that the individual losses take the following values expressed in thousand dollars: 10.1, 12.5, 14, 25, 317.3, 353, 1 200, 1 254, 52 000 and 251 000

We find that:

- $\hat{\alpha}_{\mathrm{ML}} = 15.70$ and $\hat{\beta}_{\mathrm{ML}} = 1.22$ for the log-gamma distribution
- $\hat{\alpha}_{\mathrm{ML}} = 293\,721$ and $\hat{\beta}_{\mathrm{ML}} = 0.51$ for the log-logistic distribution
- $\hat{\mu}_{\mathrm{ML}} = 12.89$ and $\hat{\sigma}_{\mathrm{ML}} = 3.35$ for the log-normal distribution
- $\hat{\mu}_{\mathrm{GMM}} = 16.26$ and $\hat{\sigma}_{\mathrm{GMM}} = 1.40$ for the log-normal distribution

An important bias

The truncation process of loss data collection

- Data are recorded only when their amounts are higher than some thresholds
- Loss thresholds vary across banks, time, business lines, etc.

"A bank must have an appropriate de minimis gross loss threshold for internal loss data collection, for example €10 000. The appropriate threshold may vary somewhat between banks, and within a bank across business lines and/or event types. However, particular thresholds should be broadly consistent with those used by peer banks" (BCBS, 2006, page 153)

Operational risk loss data

Remark

- *Operational risk loss data cannot be reduced to the sample of individual losses, but also requires specifying the threshold H_i for each individual loss x_i*
- *The form of operational loss data is then $\{(x_i, H_i), i = 1, \ldots, T\}$, where x_i is the observed value of X knowing that X is larger than the threshold H_i*

From a statistical point of view, we have:

- The true distribution is the probability distribution of X
- The sample distribution is the probability distribution of $X \mid X \geq H_i$

Dealing with loss thresholds

Analytics of the sample distribution

The sample distribution is equal to:

$$\begin{aligned}
\mathbf{F}^{\star}(x;\theta \mid H) &= \Pr\{X \leq x \mid X \geq H\} \\
&= \frac{\Pr\{X \leq x, X \geq H\}}{\Pr\{X \geq H\}} \\
&= \frac{\Pr\{X \leq x\} - \Pr\{X \leq \min(x, H)\}}{\Pr\{X \geq H\}} \\
&= \mathbb{1}\{x \geq H\} \cdot \frac{\mathbf{F}(x;\theta) - \mathbf{F}(H;\theta)}{1 - \mathbf{F}(H;\theta)}
\end{aligned}$$

It follows that the density function is:

$$f^{\star}(x;\theta \mid H) = \mathbb{1}\{x \geq H\} \cdot \frac{f(x;\theta)}{1 - \mathbf{F}(H;\theta)}$$

Dealing with loss thresholds

Analytics of the sample distribution

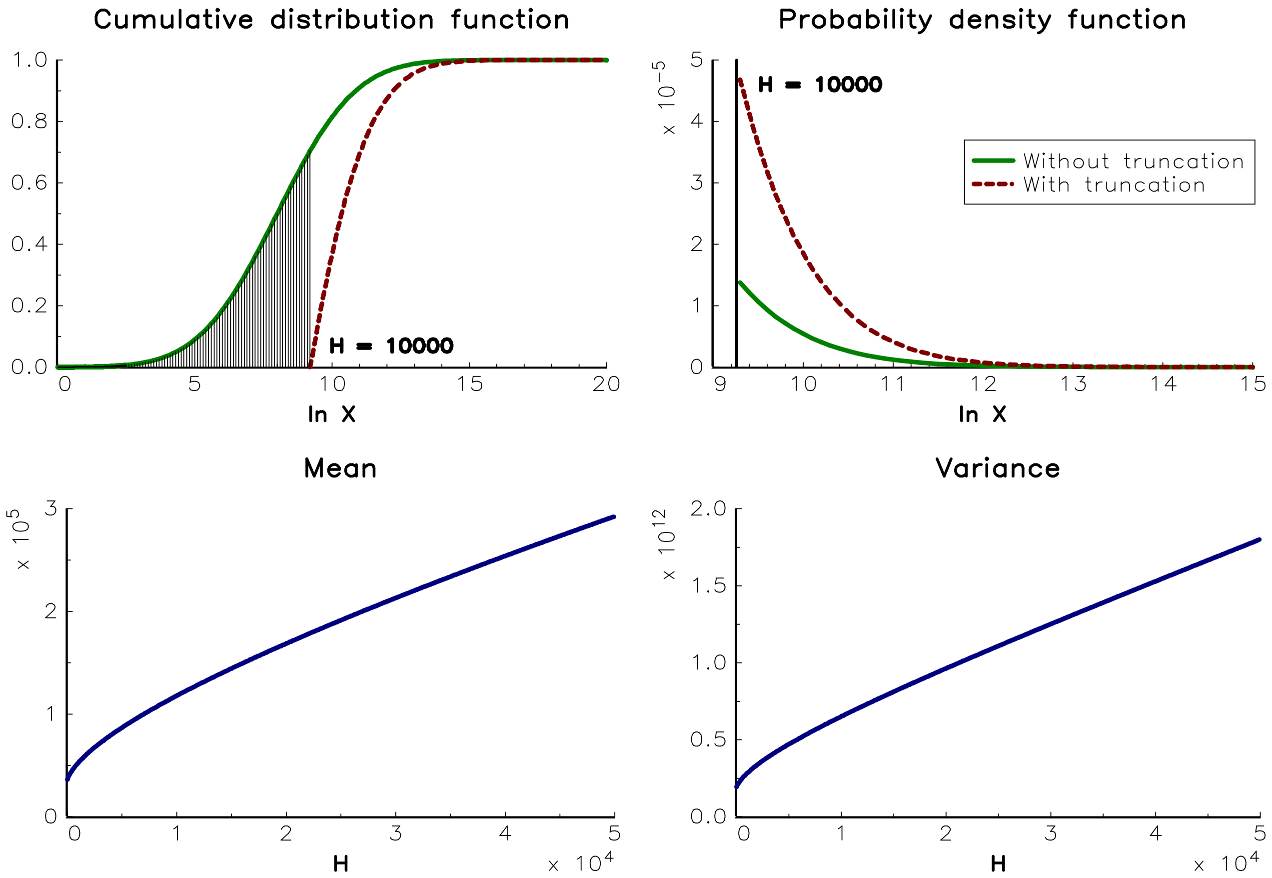


Figure: Impact of the loss threshold H on the sample distribution ($X \sim \mathcal{LN}(8, 5)$)

Dealing with loss thresholds

Application to the method of maximum likelihood

We have:

$$\begin{aligned}\ell(\theta) &= \sum_{i=1}^{T} \ln f^{\star}(x_i;\theta \mid H_i) \\ &= \sum_{i=1}^{T} \ln f(x_i;\theta) + \sum_{i=1}^{T} \ln \mathbb{1}\{x_i \geq H_i\} - \sum_{i=1}^{T} \ln(1-\mathbf{F}(H_i;\theta))\end{aligned}$$

where H_i is the threshold associated to the i^{th} observation

The correction term is $-\sum_{i=1}^{T} \ln(1-\mathbf{F}(H_i;\theta))$

Dealing with loss thresholds

Application to the method of maximum likelihood

In the case of the log-normal model, the log likelihood function is:

$$\ell(\theta) = -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln \sigma^2 - \sum_{i=1}^{T} \ln x_i - \frac{1}{2}\sum_{i=1}^{T}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2 - \sum_{i=1}^{T} \ln\left(1 - \Phi\left(\frac{\ln H_i - \mu}{\sigma}\right)\right)$$

Dealing with loss thresholds

Application to the generalized method of moments

The empirical moments become:

$$\left\{ \begin{array}{l} h_{i,1}(\theta) = x_i - \mathbb{E}[X \mid X \geq H_i] \\ h_{i,2}(\theta) = (x_i - \mathbb{E}[X \mid X \geq H_i])^2 - \mathrm{var}(X \mid X \geq H_i) \end{array} \right.$$

There is no reason that the conditional moment $\mathbb{E}[X^m \mid X \geq H_i]$ is equal to the unconditional moment $\mathbb{E}[X^m]$

Dealing with loss thresholds

Application to the generalized method of moments

In the case of the log-normal model, the empirical moments are:

$$\begin{cases} h_{i,1}(\theta) = x_i - a_1(\theta, H_i)\, e^{\mu+\frac{1}{2}\sigma^2} \\ h_{i,2}(\theta) = x_i^2 - 2x_i a_1(\theta, H_i)\, e^{\mu+\frac{1}{2}\sigma^2} + 2a_1^2(\theta, H_i)\, e^{2\mu+\sigma^2} - a_2(\theta, H_i)\, e^{2\mu+2\sigma^2} \end{cases}$$

where:

$$a_k(\theta, H) = \frac{1 - \Phi\left(\frac{\ln H - \mu - k\sigma^2}{\sigma}\right)}{1 - \Phi\left(\frac{\ln H - \mu}{\sigma}\right)}$$

Dealing with loss thresholds

Illustration

Example

We assume that the individual losses take the following values expressed in thousand dollars: 10.1, 12.5, 14, 25, 317.3, 353, 1 200, 1 254, 52 000 and 251 000

The ML estimates are $\hat{\mu}_{\text{ML}} = 12.89$ and $\hat{\sigma}_{\text{ML}} = 3.35$ for the log-normal distribution

Example

The previous losses have been collected using a unique threshold that is equal to $5 000

The ML estimates become $\hat{\mu}_{\mathrm{ML}} = 8.00$ and $\hat{\sigma}_{\mathrm{ML}} = 5.71$ for the log-normal distribution

Dealing with loss thresholds
Illustration

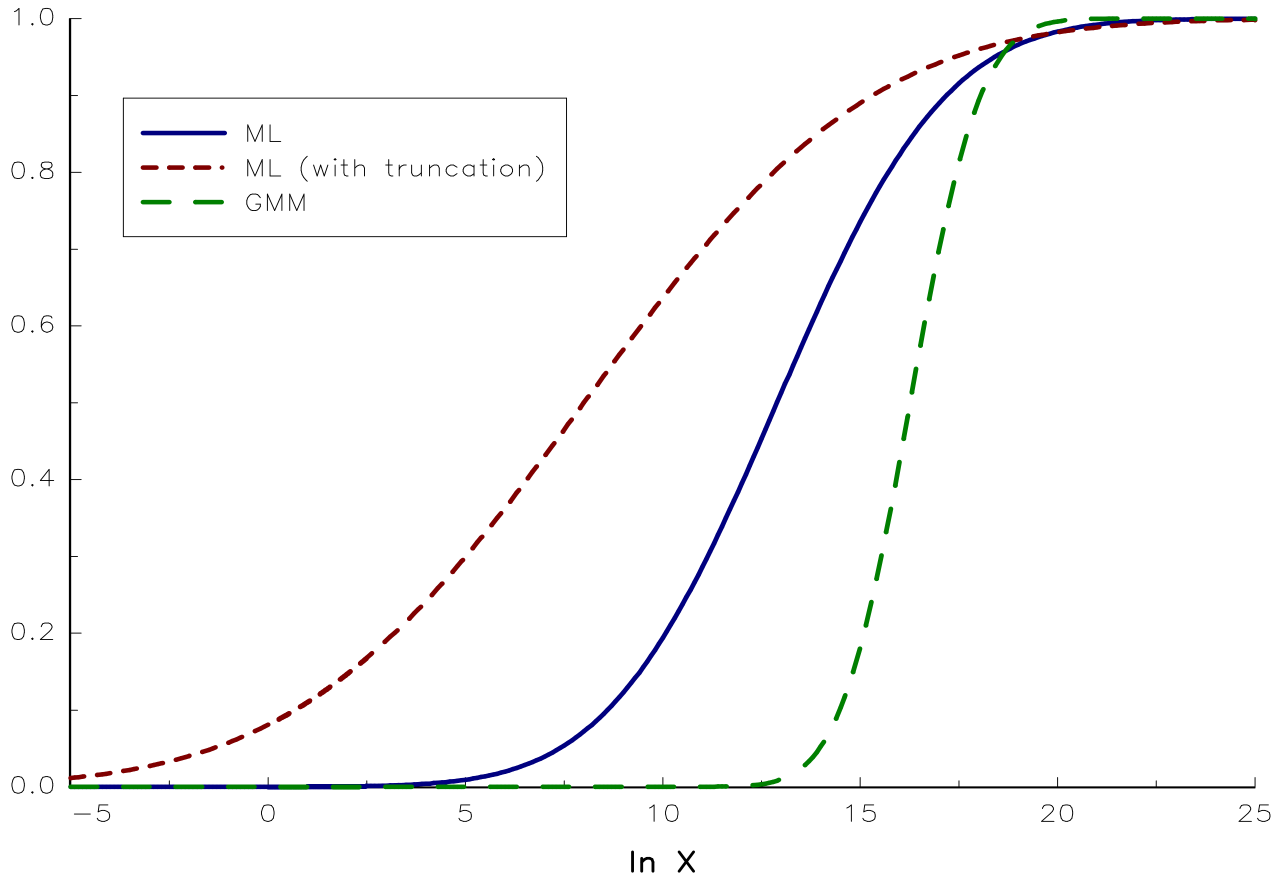


Figure: Comparison of the estimated severity distributions

Choice of the severity distribution

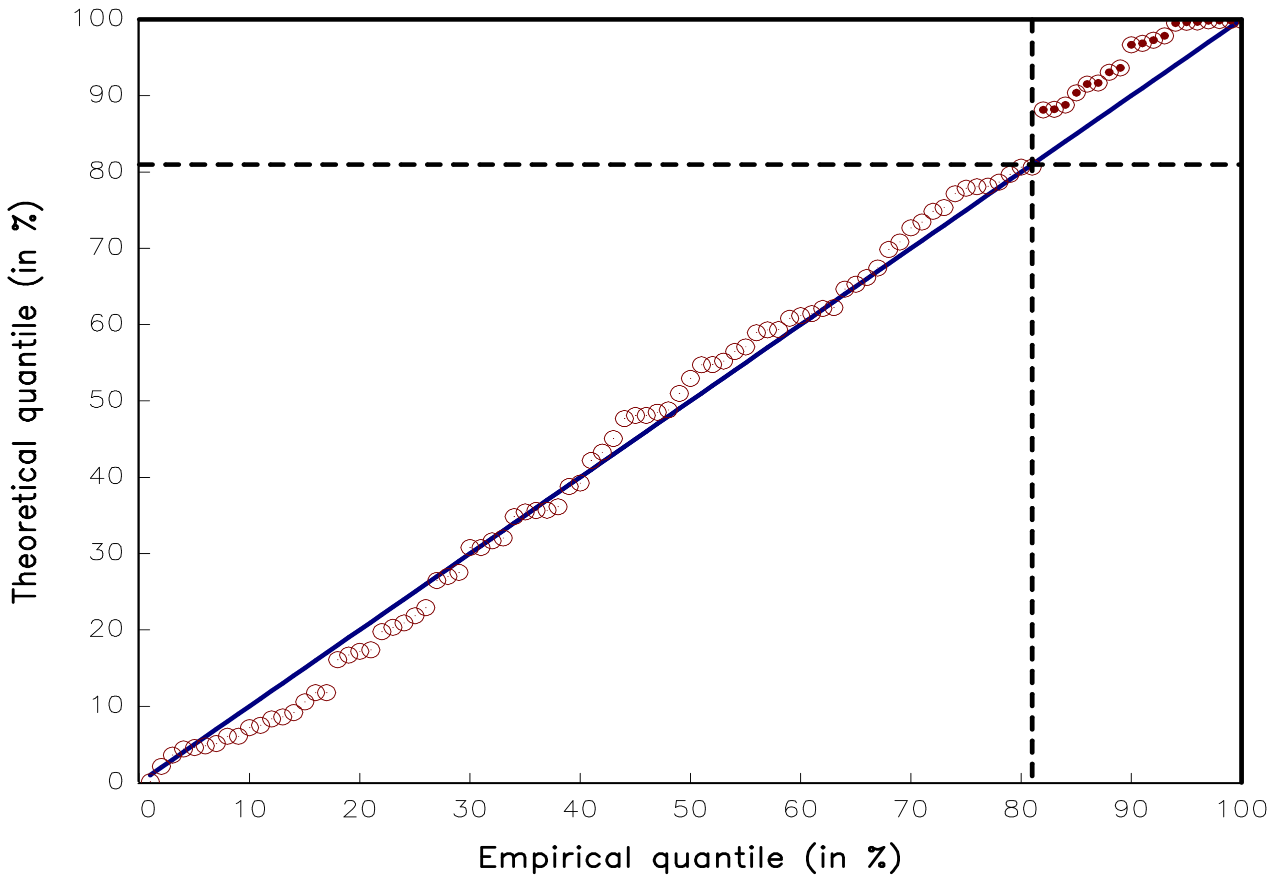


Figure: An example of QQ plot where extreme events are underestimated

Estimation of the loss frequency distribution

The goal is now to estimate P

Counting process

Let $N(t)$ be the number of losses occurring during the time period $[0, t]$. The number of losses for the time period $[t_1, t_2]$ is then equal to:

$$N(t_1; t_2) = N(t_2) - N(t_1)$$

We generally made the following statements about the process $N(t)$:

- The distribution of the number of losses $N(t; t+h)$ for each $h > 0$ is independent of t; moreover, $N(t; t+h)$ is stationary and depends only on the time interval h
- The random variables $N(t_1; t_2)$ and $N(t_3; t_4)$ are independent if the time intervals $[t_1, t_2]$ and $[t_3, t_4]$ are disjoint
- No more than one loss may occur at time t

Poisson process

These simple assumptions define a Poisson process:

1. There exists a scalar $\lambda > 0$ such that the distribution of $N(t)$ has a Poisson distribution with parameter λt
2. The duration between two successive losses is *iid* and follows the exponential distribution distribution $\mathcal{E}(\lambda)$
3. The probability mass function of the Poisson process is:

$$p(n) = \Pr\{N(t) = n\} = \frac{e^{-\lambda t} \cdot (\lambda t)^n}{n!}$$

Remark

If $N_1 \sim \mathcal{P}(\lambda_1)$ and $N_2 \sim \mathcal{P}(\lambda_2)$, then $N_1 + N_2 \sim \mathbf{P}(\lambda_1 + \lambda_2)$. We deduce that:

$$\sum_{k=1}^{K} N\left(\frac{k-1}{K}; \frac{k}{K}\right) = N(1)$$

where $N((k-1)/K; k/K) \sim \mathcal{P}(\lambda/K)$

Estimation of λ

The ML estimator is the mean of the annual number of losses:

$$\hat{\lambda}_{\mathrm{ML}} = \frac{1}{n_y} \sum_{y=1}^{n_y} N_y$$

where N_y is the number of losses occurring at year y

Since we have $\lambda = \mathbb{E}\left[N(1)\right] = \mathrm{var}\left(N(1)\right)$, the MM estimator based on the first moment is equal to:

$$\hat{\lambda}_{\mathrm{MM}} = \hat{\lambda}_{\mathrm{ML}} = \frac{1}{n_y} \sum_{y=1}^{n_y} N_y$$

whereas the MM estimator based on the first moment is equal to:

$$\hat{\lambda}_{\mathrm{MM}} = \frac{1}{n_y} \sum_{y=1}^{n_y} \left(N_y - \bar{N}\right)^2$$

where $\bar{N}$ is the average number of losses

Estimation of λ

Example

The annual number of losses from 2006 to 2015 is the following: 57, 62, 45, 24, 82, 36, 98, 75, 76 and 45. The mean is equal to 60 whereas the variance is equal to 474.40

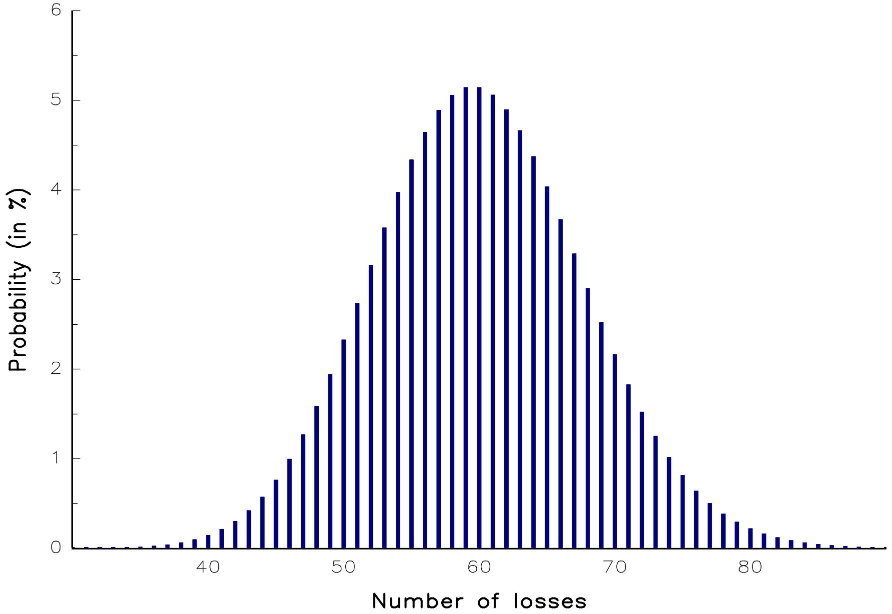


Figure: PMF of the Poisson distribution $\mathcal{P}(60)$

Not possible to observe an annual number of losses equal to 24 and 98!

Negative binomial distribution

When the variance exceeds the mean, we use the negative binomial distribution $\mathcal{NB}(r,p)$:

$$p(n) = \binom{r+n-1}{n} (1-p)^r p^n = \frac{\Gamma(r+n)}{n!\,\Gamma(r)} (1-p)^r p^n$$

where $r > 0$ and $p \in [0,1]$. We have:

$$\mathbb{E}[\mathcal{NB}(r,p)] = \frac{p \cdot r}{1-p}$$

and:

$$\mathrm{var}(\mathcal{NB}(r,p)) = \frac{p \cdot r}{(1-p)^2}$$

Remark

We verify that:

$$\mathrm{var}(\mathcal{NB}(r,p)) = \frac{1}{1-p} \cdot \mathbb{E}[\mathcal{NB}(r,p)] > \mathbb{E}[\mathcal{NB}(r,p)]$$

Negative binomial distribution and Poisson process

The negative binomial distribution corresponds to a Poisson process where the intensity parameter is random and follows a gamma distribution:

$$\begin{aligned} \mathcal{NB}(r, p) &\sim \mathcal{P}(\Lambda) \\ \Lambda &\sim \mathcal{G}(\alpha, \beta) \end{aligned}$$

where $\alpha = r$ and $\beta = (1 - p) / p$

$\Rightarrow$ See HFRM, Exercise 5.4.6, page 346 and HFRM-CB, Section 5.4.6, pages 113-116

Application to the example

- Using the previous example, we obtain:

$$\hat{r}_{\mathrm{MM}} = \frac{m^2}{v - m} = \frac{60^2}{474.40 - 60} = 8.6873$$

and

$$\hat{p}_{\mathrm{MM}} = \frac{v - m}{v} = \frac{474.40 - 60}{474.40} = 0.8735$$

where m is the mean and v is the variance of the sample

- If we use the method of maximum likelihood, we obtain $\hat{r}_{\mathrm{ML}} = 7.7788$ and $\hat{p}_{\mathrm{ML}} = 0.8852$
- We deduce that:

$$\mathcal{NB}\left(\hat{r}_{\mathrm{ML}}, \hat{p}_{\mathrm{ML}}\right) \sim \mathcal{P}\left(\Lambda\right)$$

where:

$$\Lambda \sim \mathcal{G}\left(7.7788, 0.1296\right)$$

- $\mathcal{P}\left(60\right)$ and $\mathcal{NB}\left(\hat{r}_{\mathrm{ML}}, \hat{p}_{\mathrm{ML}}\right)$ have the same mean, but not the same variance

Application to the example

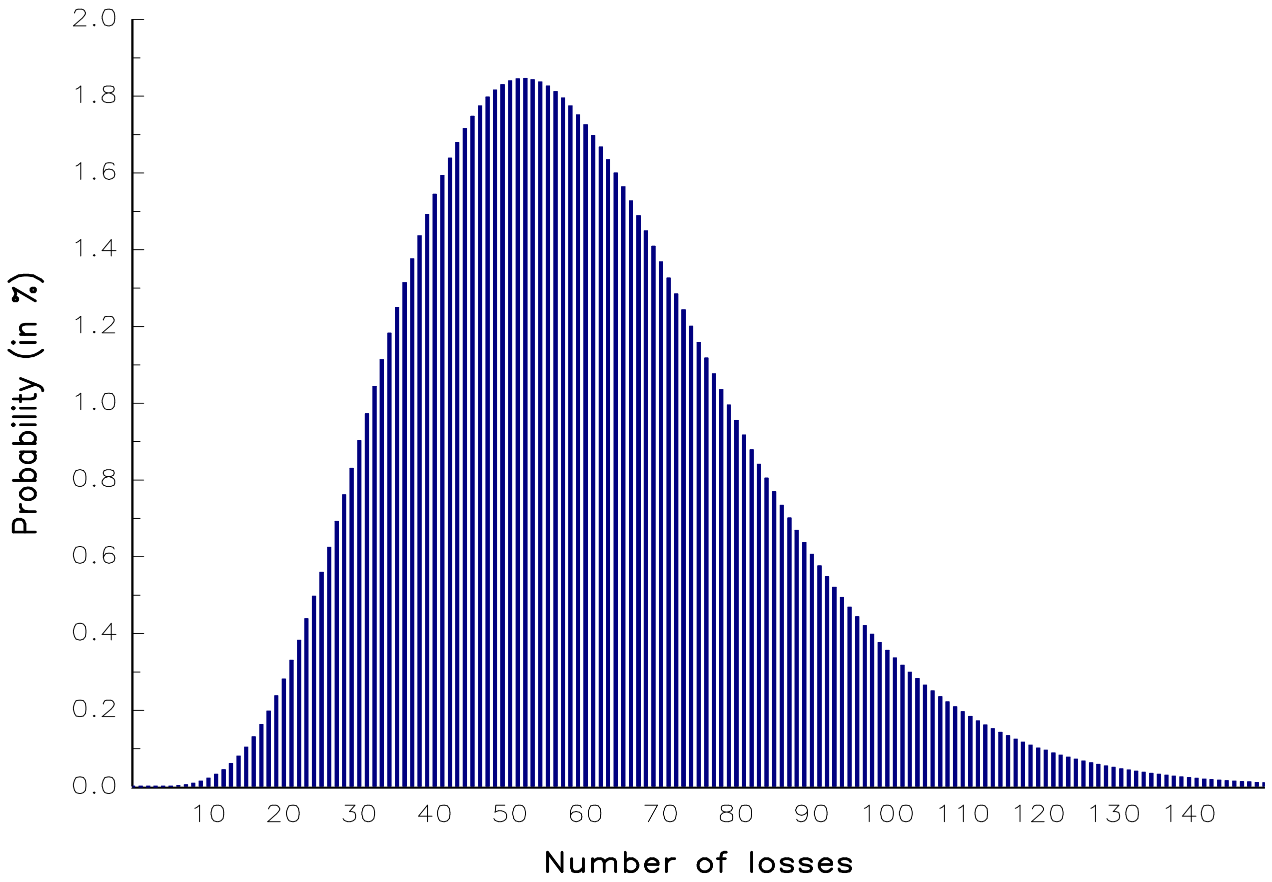


Figure: PMF of the negative binomial distribution

Application to the example

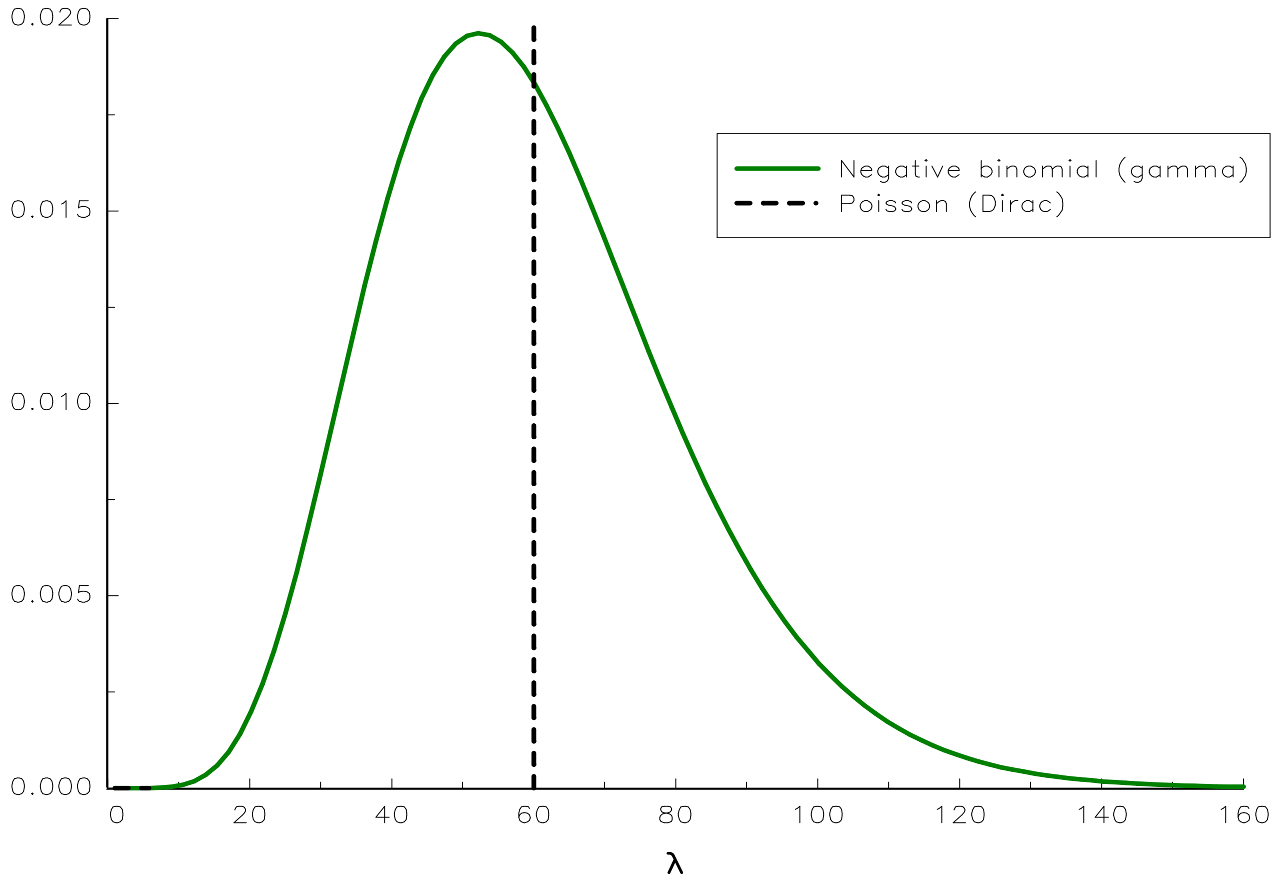


Figure: Probability density function of the parameter λ

Dealing with a loss threshold

- The loss threshold has an impact on the sample frequency distribution
- For instance, if the threshold H is set at a high level, then the average number of reported losses is low
- Let $N_H(t)$ be the number of events that are larger than the threshold H:

$$N_H(t) = \sum_{i=1}^{N(t)} \mathbb{1}\{X_i > H\}$$

- We can show that (HFRM, page 326):

$$\mathbb{E}[N_H(t)] = \mathbb{E}[N(t)] \cdot \Pr\{X_i > H\} = \mathbb{E}[N(t)] \cdot (1 - \mathbf{F}(H;\theta))$$

Dealing with a loss threshold

In the case of the Poisson distribution, we also prove that:

$$\mathbf{P}_H(\lambda) = \mathbf{P}(\lambda_H)$$

We deduce that the estimator $\hat{\lambda}$ has the following expression:

$$\hat{\lambda} = \frac{\hat{\lambda}_H}{1 - \mathbf{F}\left(H; \hat{\theta}\right)}$$

where:

- $\hat{\lambda}_H$ is the average number of losses that are collected above the threshold H
- $\mathbf{F}\left(x; \hat{\theta}\right)$ is the parametric estimate of the severity distribution.

Remark

This approach is only valid if the loss threshold is unique

Dealing with a loss threshold

Example

We consider that the bank has collected the loss data from 2006 to 2015 with a threshold of \$20 000. For a given event type, the calibrated severity distribution corresponds to a log-normal distribution with parameters $\hat{\mu} = 7.3$ and $\hat{\sigma} = 2.1$, whereas the annual number of losses is the following: 23, 13, 50, 12, 25, 36, 48, 27, 18 and 35

Using the Poisson distribution, we obtain $\hat{\lambda}_H = 28.70$. The probability that the loss exceeds the threshold H is equal to:

$$\Pr \{X > 20\,000\} = 1 - \Phi \left(\frac{\ln (20\,000) - 7.3}{2.1} \right) = 10.75\%$$

This means that only 10.75% of losses can be observed when we apply a threshold of \$20 000. We deduce that the estimate of the Poisson parameter is equal to:

$$\hat{\lambda} = \frac{28.70}{10.75\%} = 266.90$$

Calculating the capital charge

Several approaches:

- Monte Carlo approach
- Method of characteristic functions
- Panjer recursive approach
- Single loss approximation

Monte Carlo approach

Algorithm

Compute the capital-at-risk for an operational risk cell

Initialize the number of simulations n_S
for $j = 1 : n_S$ **do**
  Simulate an annual number n of losses from the frequency distribution $\mathbf{P}$
  $S_j \leftarrow 0$
  for $i = 1 : n$ **do**
    Simulate a loss X_i from the severity distribution $\mathbf{F}$
    $S_j = S_j + X_i$
  end for
end for
Calculate the order statistics $S_{1:n_S}, \ldots, S_{n_S:n_S}$
Deduce the capital-at-risk $\mathrm{CaR} = S_{\alpha n_S:n_S}$ with $\alpha = 99.9\%$
return CaR

Monte Carlo approach

Illustration

- We assume that $N(1) \sim \mathcal{P}(4)$ and $X_i \sim \mathcal{LN}(8,4)$
- The simulated values of $N(1)$ are 3, 4, 1, 2, 3, etc.
- The simulated values of X_i are 3388.6, 259.8, 13328.3, 39.7, 1220.8, 1486.4, 15197.1, 3205.3, 5070.4, 84704.1, 64.9, 1237.5, 187073.6, 4757.8, 50.3, 2805.7, etc.

For the first simulation, we have three losses and we obtain:

$$S_1 = 3388.6 + 259.8 + 13328.3 = \$16\,976.7$$

For the second simulation, the number of losses is equal to four and the compound loss is equal to:

$$S_2 = 39.7 + 1220.8 + 1486.4 + 15197.1 = \$17\,944.0$$

For the third simulation, we obtain:

$$S_3 = \$3\,205.3$$

Monte Carlo approach

The Monte Carlo method is powerful and the most used approach for computing the capital charge for operational risk

But be careful about the convergence!

Panjer recursion

Theorem

Panjer (1981) showed that if the pmf of the counting process $N(t)$ satisfies:

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1)$$

where a and b are two scalars, then the following recursion holds:

$$g(x) = p(1) f(x) + \int_0^x \left(a + b\frac{y}{x}\right) f(y) g(x-y) \, \mathrm{d}y$$

where $x > 0$

Panjer recursion

For discrete severity distributions satisfying $f_n = \Pr\{X_i = n\delta\}$ where δ is the monetary unit (e.g. \$10 000), the Panjer recursion becomes:

$$g_n = \Pr\{S = n\delta\} = \frac{1}{1 - af_0} \sum_{j=1}^{n} \left(a + \frac{bj}{n}\right) f_j g_{n-j}$$

where:

$$g_0 = \sum_{n=0}^{\infty} p(n) (f_0)^n = \begin{cases} p(0) e^{bf_0} & \text{if } a = 0 \\ p(0) (1 - af_0)^{-1-b/a} & \text{otherwise} \end{cases}$$

The capital-at-risk is then equal to:

$$\mathrm{CaR}(\alpha) = n^\star \delta$$

where:

$$n^\star = \inf \left\{ n : \sum_{j=0}^{n} g_j \geq \alpha \right\}$$

Panjer recursion

Example

We consider the compound Poisson distribution with log-normal losses and different sets of parameters:

(a) $\lambda = 5$, $\mu = 5$, $\sigma = 1.0$
(b) $\lambda = 5$, $\mu = 5$, $\sigma = 1.5$
(c) $\lambda = 5$, $\mu = 5$, $\sigma = 2.0$
(d) $\lambda = 50$, $\mu = 5$, $\sigma = 2.0$

We perform a discretization of the severity distribution:

$$f_n = \Pr\left\{ n\delta - \frac{\delta}{2} \leq X_i \leq n\delta + \frac{\delta}{2} \right\} = \mathbf{F}\left(n\delta + \frac{\delta}{2} \right) - \mathbf{F}\left(n\delta - \frac{\delta}{2} \right)$$

with the convention $f_0 = \mathbf{F}(\delta/2)$

Panjer recursion

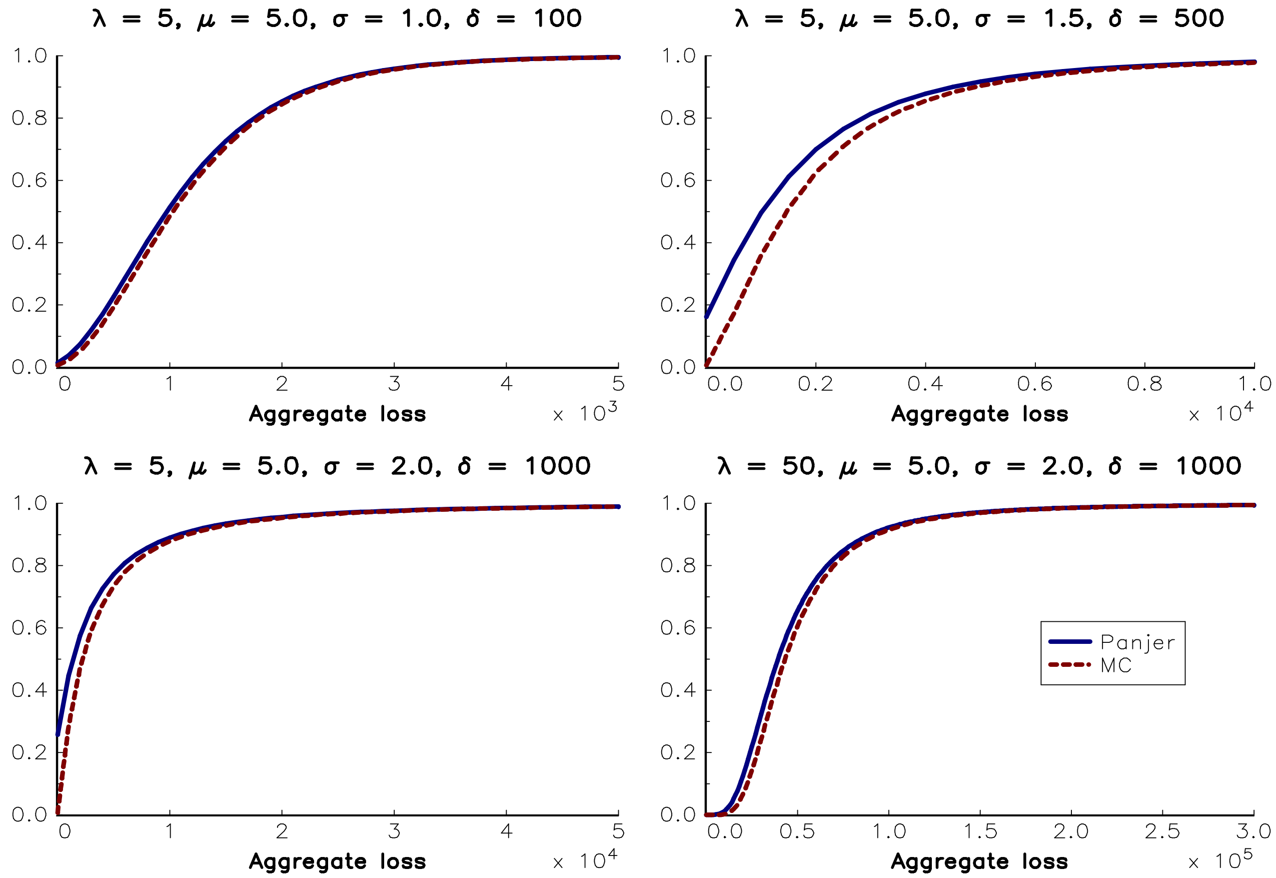


Figure: Comparison between the Panjer and MC compound distributions

Panjer recursion

Table: Comparison of the capital-at-risk calculated with Panjer recursion and Monte Carlo simulations

α	Panjer recursion				Monte Carlo simulations			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
90%	2400	4500	11000	91000	2350	4908	11648	93677
95%	2900	6500	19000	120000	2896	6913	19063	123569
99%	4300	13500	52000	231000	4274	13711	51908	233567
99.5%	4900	18000	77000	308000	4958	17844	77754	310172
99.9%	6800	32500	182000	604000	6773	32574	185950	604756

Single loss approximation

If the severity belongs to the family of subexponential distributions (HFRM, pages 333-336), Böcker and Klüppelberg (2005) showed that:

$$\begin{aligned}
\mathbf{G}^{-1}(\alpha) &= \mathrm{EL} + \mathrm{UL}(\alpha) \\
&\approx \mathbb{E}[N(1)] \cdot \mathbb{E}[X_i] + \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{N(1)}\right) - \mathbb{E}[X_i] \\
&\approx (\mathbb{E}[N(1)] - 1) \cdot \mathbb{E}[X_i] + \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{\mathbb{E}[N(1)]}\right)
\end{aligned}$$

If $N(1) \sim \mathcal{P}(\lambda)$ and $X_i \sim \mathcal{LN}(\mu, \sigma^2)$, we obtain:

$$\mathrm{EL} = \lambda \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

and:

$$\mathrm{UL}(\alpha) \approx \exp\left(\mu + \sigma\Phi^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right)\right) - \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

Single loss approximation

Remark

A better approximation of the capital-at-risk is:

$$\mathbf{G}^{-1}(\alpha) \approx \left(\mathbf{P}^{-1}(\alpha) - 1\right)\mathbb{E}[X_i] + \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{\mathbb{E}[N(1)]}\right)$$

where $\mathbf{P}$ *is the cumulative distribution function of the counting process* $N(1)$

How to compute the total capital charge?

The operational risk loss L of the bank is divided into a matrix of homogenous losses:

$$L = \sum_{k=1}^{K} S_k$$

where S_k is the sum of losses of the k^{th} cell and K is the number of cells in the matrix (Basel II $= 7 \times 8 = 56$ cells)

Using LDA, we know how to compute S_k. **But how to compute the total loss L?**

The solution is given by the copula approach

It only works with the Monte Carlo approach and uses the method of the empirical quantile function (HFRM, Section 13.1.3.2, pages 805-808)

Probability distribution of a given scenario

We assume that $N(t) \sim \mathcal{P}(\lambda)$. Let τ_n be the arrival time of the n^{th} loss:

$$\tau_n = \inf \{t \geq 0 : N(t) = n\}$$

- We know that $T_n = \tau_n - \tau_{n-1} \sim \mathcal{E}(\lambda)$
- We recall that the losses $X_n \sim \mathbf{F}$
- We note $T_n(x)$ the duration between two losses exceeding x
- We have $T_n(x) \equiv T_1(x)$

Theorem

We have:

$$T_n(x) \sim \mathcal{E}(\lambda(1 - \mathbf{F}(x)))$$

and:

$$\mathbb{E}[T_n(x)] = \frac{1}{\lambda(1 - \mathbf{F}(x))}$$

Probability distribution of a given scenario

Proof

By using the fact that a finite sum of exponential times is an Erlang distribution, we have:

$$\begin{aligned}
\Pr\left\{T_1(x) > t\right\} &= \sum_{n\geq 1} \Pr\left\{\tau_n > t; X_1 < x, \ldots, X_{n-1} < x; X_n \geq x\right\} \\
&= \sum_{n\geq 1} \Pr\left\{\tau_n > t\right\} \cdot \mathbf{F}(x)^{n-1} \cdot (1 - \mathbf{F}(x)) \\
&= \sum_{n\geq 1} \mathbf{F}(x)^{n-1} \cdot (1 - \mathbf{F}(x)) \cdot \left(\sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}\right) \\
&= (1 - \mathbf{F}(x)) \cdot \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \left(\sum_{n=k}^{\infty} \mathbf{F}(x)^n\right) \\
&= e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \mathbf{F}(x)^k \\
&= e^{-\lambda(1-\mathbf{F}(x))t}
\end{aligned}$$

Calibration of a set of scenarios

- We define a scenario as "a loss of x or higher occurs once every d years"
- We assume that the severity distribution is $\mathbf{F}(x;\theta)$ and the frequency distribution is $\mathcal{P}(\lambda)$
- Suppose that we face different scenarios $\{(x_s, d_s), s = 1, \ldots, n_S\}$. We may estimate the implied parameters underlying the expert judgements using the method of moments:

$$\left(\hat{\lambda}_{\mathrm{MM}}, \hat{\theta}_{\mathrm{MM}}\right) = \arg\min \sum_{s=1}^{n_S} w_s \cdot \left(d_s - \frac{1}{\lambda\left(1 - \mathbf{F}(x_s;\theta)\right)}\right)^2$$

  where w_s is the weight of the s^{th} scenario
- We can show that the optimal weights w_s correspond to the inverse of the variance of d_s:

$$w_s = \frac{1}{\mathrm{var}(d_s)} = \lambda\left(1 - \mathbf{F}(x_s;\theta)\right)$$

Calibration of a set of scenarios

Numerical solution

To solve the previous optimization program, we proceed by iterations:

- Let $\left(\hat{\lambda}_m, \hat{\theta}_m\right)$ be the solution of the minimization program:

$$\left(\hat{\lambda}_m, \hat{\theta}_m\right) = \arg\min \sum_{j=1}^{p} \hat{\lambda}_{m-1} \cdot \left(1 - \mathbf{F}\left(x_s; \hat{\theta}_{m-1}\right)\right) \cdot \left(d_s - \frac{1}{\lambda\left(1 - \mathbf{F}\left(x_s; \theta\right)\right)}\right)^2$$

- Under some conditions, the estimator $\left(\hat{\lambda}_m, \hat{\theta}_m\right)$ converge to the optimal solution
- We can simplify the optimization program by using the following approximation:

$$w_s = \frac{1}{\mathrm{var}\left(d_s\right)} = \frac{1}{\mathbb{E}\left[d_s\right]} \simeq \frac{1}{d_s}$$

Calibration of a set of scenarios

Example

We assume that the severity distribution is log-normal and consider the following set of expert's scenarios:

x_s (in \$ mn)	1	2.5	5	7.5	10	20
d_s (in years)	1/4	1	3	6	10	40

Calibration of a set of scenarios

#1 If $w_s = 1$, we obtain $\hat{\lambda} = 43.400$, $\hat{\mu} = 11.389$ and $\hat{\sigma} = 1.668$

#2 Using the approximation $w_s \simeq 1/d_s$, the estimates become $\hat{\lambda} = 154.988$, $\hat{\mu} = 10.141$ and $\hat{\sigma} = 1.855$

#3 The optimal estimates are $\hat{\lambda} = 148.756$, $\hat{\mu} = 10.181$ and $\hat{\sigma} = 1.849$

Here are the estimated values of the duration:

x_s (in \$ mn)	1	2.5	5	7.5	10	20
#1	0.316	1.022	2.964	5.941	10.054	39.997
#2	0.271	0.968	2.939	5.973	10.149	39.943
#3	0.272	0.970	2.941	5.974	10.149	39.944

Exercises

- Severity distribution
  - Exercise 5.4.1 – Estimating the loss severity distribution
  - Exercise 5.4.5 – Parametric estimation of the loss severity distribution
- Frequency distribution
  - Exercise 5.4.2 – Estimation of the loss frequency distribution
- Other topics
  - Exercise 5.4.3 – Using the method of moments in operational risk models
  - Exercise 5.4.6 – Mixed Poisson process

References

- Basel Committee on Banking Supervision (2006)
  International Convergence of Capital Measurement and Capital Standards — A Revised Framework — Comprehensive version, June 2006.
- Basel Committee on Banking Supervision (2017)
  Basel III: Finalising Post-crisis Reforms, December 2017.
- Roncalli, T. (2020)
  Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 5.
- Roncalli, T. (2020)
  Handbook of Financial Risk Management – Companion Book, Chapter 5.

Course 2023-2024 in Financial Risk Management
Lecture 6. Liquidity Risk

Thierry Roncalli*

*Amundi Asset Management[15]

*University of Paris-Saclay

September 2023

[15]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

Bid-ask spread

Definition

The bid-ask quoted spread $\mathbf{S}_t$ is defined by:

$$\mathbf{S}_t = \frac{P_t^{\mathrm{ask}} - P_t^{\mathrm{bid}}}{P_t^{\mathrm{mid}}}$$

where P_t^{ask}, P_t^{bid} and P_t^{mid} are the ask, bid and mid quotes for a given security at time t.

We have:

$$P_t^{\mathrm{mid}} = \frac{P_t^{\mathrm{ask}} + P_t^{\mathrm{bid}}}{2}$$

Bid-ask spread

Table: Snapshot of the limit order book of the Lyxor Euro Stoxx 50 ETF recorded at NYSE Euronext Paris – The corresponding date is 14:00:00 and $56,566$ micro seconds on 28 December 2012

i^{th} limit	Buy orders		Sell orders	
	$Q_t^{\text{bid,i}}$	$P_t^{\text{bid,i}}$	$Q_t^{\text{ask,i}}$	$P_t^{\text{ask,i}}$
1	65 201	26.325	70 201	26.340
2	85 201	26.320	116 201	26.345
3	105 201	26.315	107 365	26.350
4	76 500	26.310	35 000	26.355
5	20 000	26.305	35 178	26.360

We have $P_t^{\text{bid}} = 26.325$ and $P_t^{\text{ask}} = 26.340$, implying that the mid price is equal to $P_t^{\text{mid}} = (26.325 + 26.340)/2 = 26.3325$. We deduce that the bid-ask spread is:

$$\mathbf{s}_t = \frac{26.340 - 26.325}{26.3325} = 5.696 \text{ bps}$$

Bid-ask spread

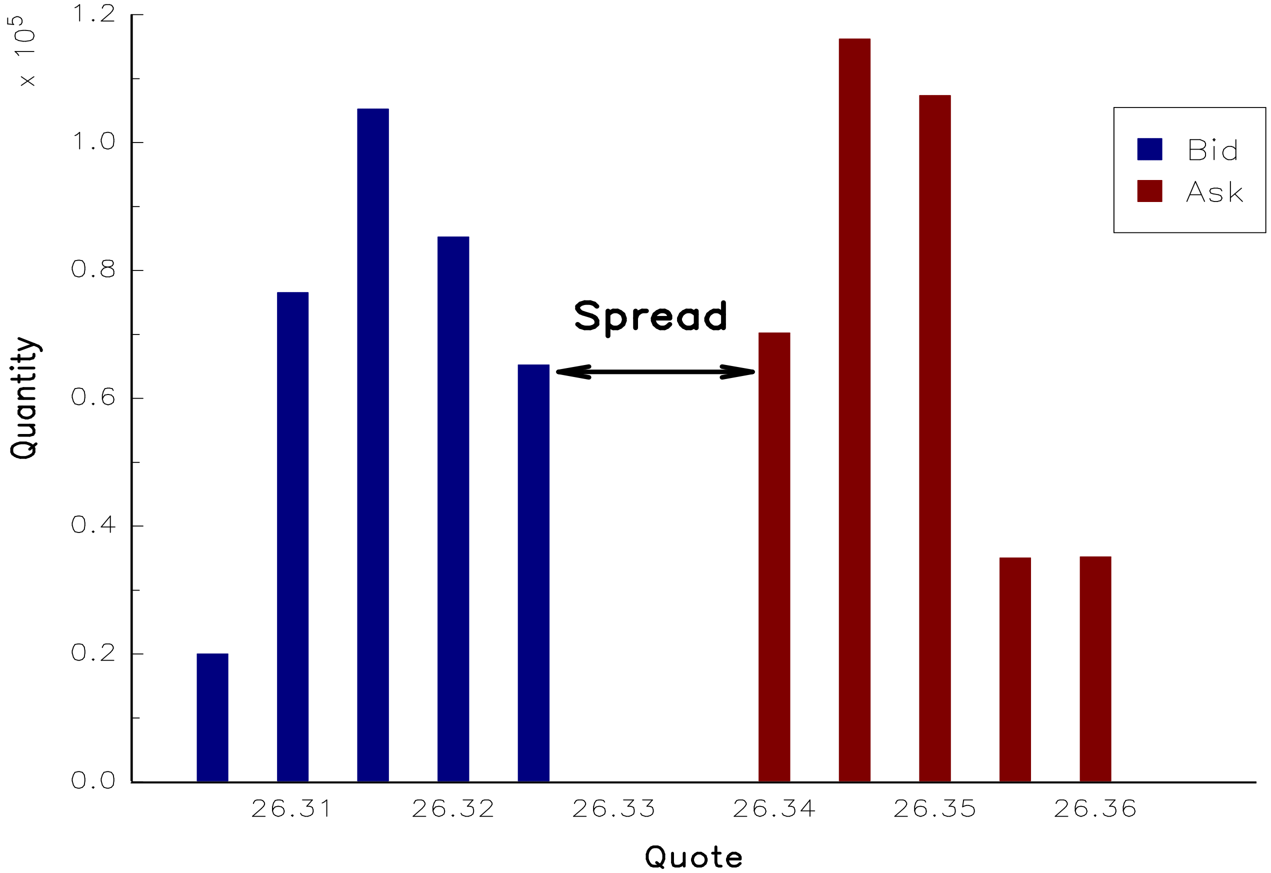


Figure: An example of a limit order book

Bid-ask spread

- The *effective spread* is equal to:

$$\mathbf{S}_{\tau}^{e} = 2 \left| \frac{P_{\tau} - P_{t}^{\mathrm{mid}}}{P_{t}^{\mathrm{mid}}} \right|$$

where τ is the trade index, P_{τ} is the price of the τ^{th} trade and P_{τ}^{mid} is the midpoint of market quote calculated at the time t of the τ^{th} trade

- The *realized spread* is equal to:

$$\mathbf{S}_{\tau}^{r} = 2 \left| \frac{P_{\tau} - P_{t+\Delta}^{\mathrm{mid}}}{P_{t+\Delta}^{\mathrm{mid}}} \right|$$

Generally, Δ is set to five minutes

Price impact $\Rightarrow P_{t+\Delta}^{\mathrm{mid}} \neq P_{t}^{\mathrm{mid}}$

Trading volume

The trading volume $\mathbf{V}_t$ indicates the dollar value of the security exchanged during the period t:

$$\mathbf{V}_t = \sum_{\tau \in t} Q_\tau P_\tau$$

where Q_τ and P_τ are the τ^{th} quantity and price traded during the period. Generally, we consider a one-day period and use the following approximation:

$$\mathbf{V}_t \approx Q_t P_t$$

where Q_t is the number of securities traded during the day t and P_t is the closing price of the security.

Turnover

The turnover is the ratio between the trading volume and the free float market capitalization M_t of the asset:

$$\mathbf{T}_t = \frac{\mathbf{V}_t}{M_t} = \frac{\mathbf{V}_t}{N_t P_t}$$

where N_t is the number of outstanding '*floating*' shares

$\Rightarrow$ The asset turnover ratio indicates how many times each share changes hands in a given period

Liquidation ratio

The liquidation ratio $\mathcal{LR}(m)$ measures the proportion of a given position that can be liquidated after m trading days

Liquidation ratio

Computation of the liquidation ratio

We denote $(x_1, \ldots, x_n)$ the number of shares held in the portfolio. For each asset that composes the portfolio, we denote x_i^+ the maximum number of shares for asset i that can be sold during a trading day. The number of shares $x_i(m)$ liquidated after m trading days is defined as follows:

$$x_i(m) = \min\left(\left(x_i - \sum_{k=0}^{m-1} x_i(k)\right)^+, x_i^+\right)$$

with $x_i(0) = 0$. The liquidation ratio $\mathcal{LR}(m)$ is then the proportion of the portfolio liquidated after m trading days:

$$\mathcal{LR}(m) = \frac{\sum_{i=1}^n \sum_{k=0}^m x_i(k) \cdot P_{i,t}}{\sum_{i=1}^n x_i \cdot P_{i,t}}$$

Liquidation ratio

Table: Statistics of the liquidation ratio (size = \$10 bn, liquidation policy = 10% of ADV)

Statistics	SPX	SX5E	DAX	NDX	MSCI EM	MSCI INDIA	MSCI EMU SC
m (in days)	Liquidation ratio $\mathcal{LR}(t)$ in %						
1	88.4	12.3	4.8	40.1	22.1	1.5	3.0
2	99.5	24.7	9.6	72.6	40.6	3.0	6.0
5	100.0	58.8	24.1	99.7	75.9	7.6	14.9
10	100.0	90.1	47.6	99.9	93.9	15.1	29.0
α (in %)	Liquidation time $\mathcal{LR}^{-1}(\alpha)$ in days						
50	1	5	11	2	3	37	21
75	1	7	17	3	5	71	43
90	2	10	23	3	9	110	74
99	2	15	29	5	17	156	455

Source: Roncalli and Weisang (2015).

Liquidation ratio

Table: Statistics of the liquidation ratio (size = \$10 bn, liquidation policy = 30% of ADV)

Statistics	SPX	SX5E	DAX	NDX	MSCI EM	MSCI INDIA	MSCI EMU SC
t (in days)	Liquidation ratio $\mathcal{LR}(t)$ in %						
1	100.0	37.0	14.5	91.0	55.5	4.5	9.0
2	100.0	67.7	28.9	99.8	81.8	9.1	17.8
5	100.0	99.2	68.6	100.0	98.5	22.6	40.4
10	100.0	100.0	99.6	100.0	100.0	43.1	63.2
α (in %)	Liquidation time $\mathcal{LR}^{-1}(\alpha)$ in days						
50	1	2	4	1	1	13	7
75	1	3	6	1	2	24	15
90	1	4	8	1	3	37	25
99	1	5	10	2	6	52	152

Source: Roncalli and Weisang (2015).

Other liquidity measures

- Hui-Heubel liquidity ratio

$$\mathbf{H}_t^2 = \frac{1}{\mathbf{T}_t} \left(\frac{P_t^{\mathrm{high}} - P_t^{\mathrm{low}}}{P_t^{\mathrm{low}}} \right)$$

- Hasbrouck-Schwartz variance ratio

$$\mathbf{VR} = \frac{\mathrm{var}\left(R_{t,t+h}\right)}{\mathrm{var}\left(R_{t,t+1}\right)}$$

- Amihud measure

$$\mathbf{ILLIQ} = \frac{1}{n_t} \sum_t \frac{\left|R_{t,t+1}\right|}{\mathbf{V}_t}$$

- Implicit spread of Roll (1984):

$$\tilde{\mathbf{S}} = 2\sqrt{-\operatorname{cov}\left(\Delta P_t, \Delta P_{t-1}\right)}$$

L-CAPM

"[...] there is also broad belief among users of financial liquidity – traders, investors and central bankers – that the principal challenge is not the average level of financial liquidity... but its variability and uncertainty" (Persaud, 2003).

L-CAPM

We note $R_{i,t}$ and $L_{i,t}$ the gross return and the relative (stochastic) illiquidity cost of Asset i. At the equilibrium, Acharya and Pedersen (2005) showed that:

$$\mathbb{E}\left[R_{i,t} - L_{i,t}\right] - r = \tilde{\beta}_i \cdot \left(\mathbb{E}\left[R_{m,t} - L_{m,t}\right] - r\right)$$

where r is the return of the risk-free asset, $R_{m,t}$ and $L_{m,t}$ are the gross return and the illiquidity cost of the market portfolio, and $\tilde{\beta}_i$ is the liquidity-adjusted beta of Asset i:

$$\tilde{\beta}_i = \frac{\operatorname{cov}\left(R_{i,t} - L_{i,t}, R_{m,t} - L_{m,t}\right)}{\operatorname{var}\left(R_{m,t} - L_{m,t}\right)}$$

Asset liability mismatch

"We define funding liquidity as the ability to settle obligations with immediacy. Consequently, a bank is illiquid if it is unable to settle obligations. Legally, a bank is then in default. Given this definition we define funding liquidity risk as the possibility that over a specific horizon the bank will become unable to settle obligations with immediacy" (Drehmann and Nikolaou, 2013).

Relationship between market and funding liquidity risks

"Traders provide market liquidity, and their ability to do so depends on their availability of funding. Conversely, traders' funding, i.e., their capital and margin requirements, depends on the assets' market liquidity. We show that, under certain conditions, margins are destabilizing and market liquidity and funding liquidity are mutually reinforcing, leading to liquidity spirals" (Brunnermeier and Pedersen, 2009).

Relationship between market and funding liquidity risks

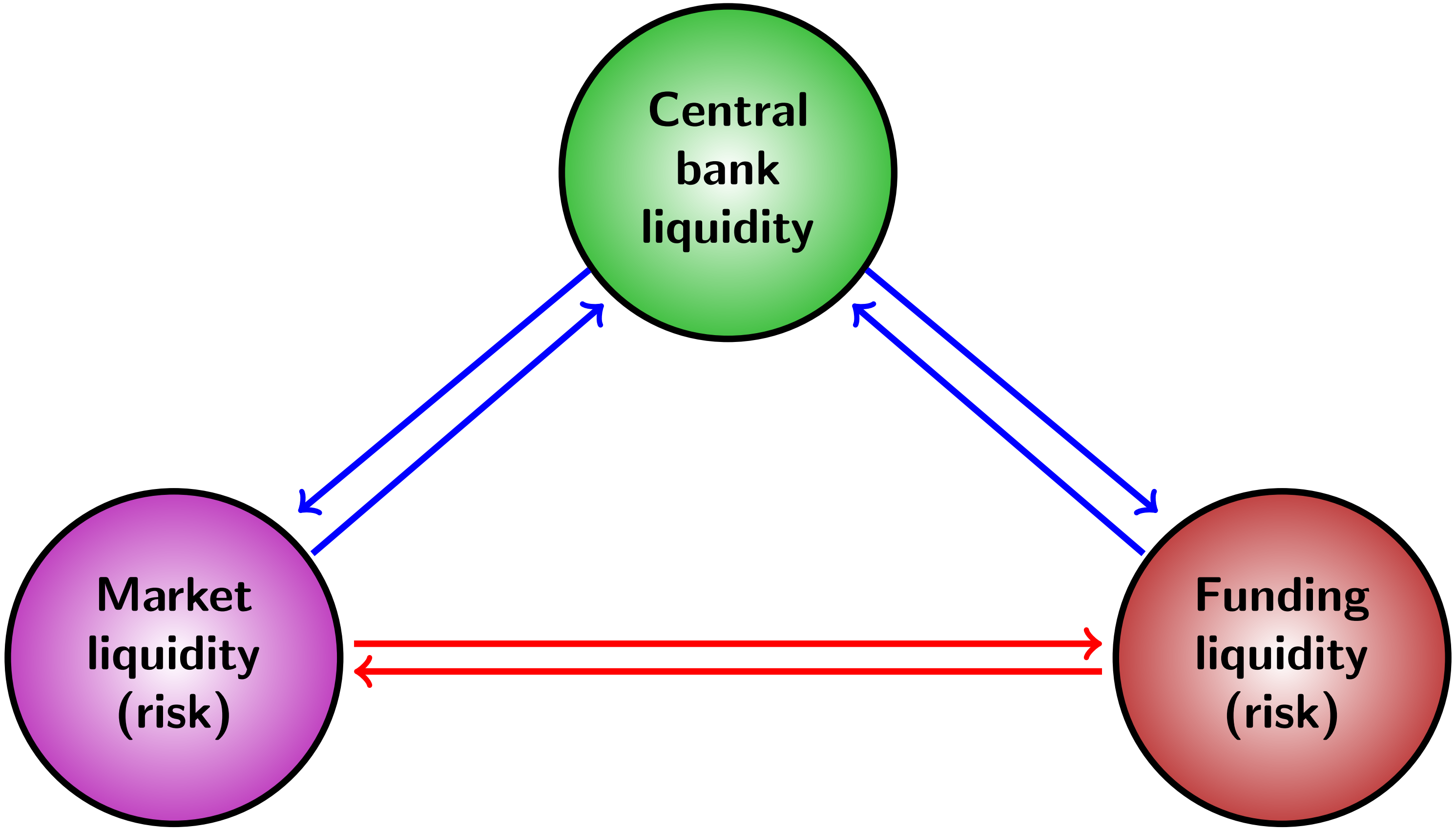


Figure: The liquidity nodes of the financial system

Source: Nikolaou (2009).

Relationship between market and funding liquidity risks

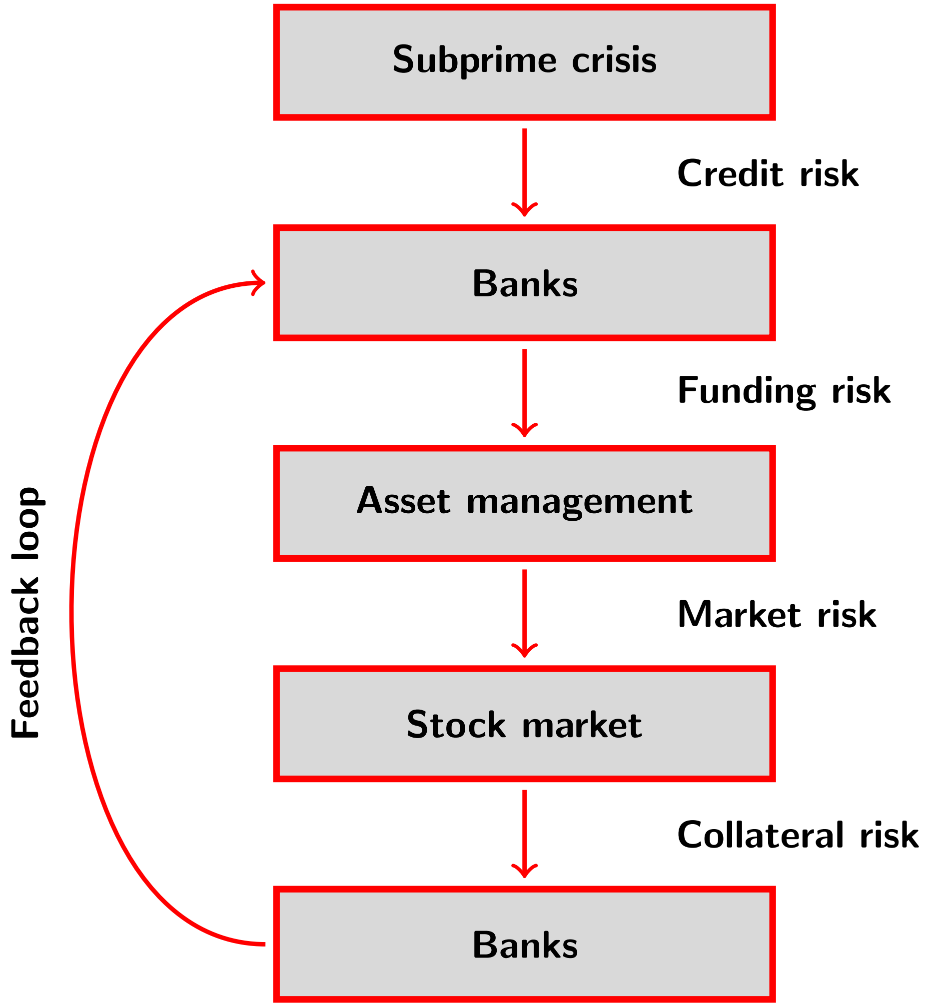


Figure: Spillover effects during the 2008 global financial crisis

Liquidity coverage ratio

The liquidity coverage ratio is defined as:

$$\mathrm{LCR} = \frac{\text{HQLA}}{\text{Total net cash outflows}} \geq 100\%$$

where the numerator is the stock of high quality liquid assets (HQLA) in stressed conditions, and the denominator is the total net cash outflows over the next 30 calendar days

$\Rightarrow$ The underlying idea of the LCR is that the bank has sufficient liquid assets to meet its liquidity needs for the next month

High quality liquid asset

An asset is considered to be a HQLA if it can be easily converted into cash. Therefore, the concept of HQLA is related to asset quality and asset liquidity

Characteristics used by the Basel Committee for defining HQLA:

- fundamental characteristics (low risk, ease and certainty of valuation, low correlation with risky assets, listed on a developed and recognized exchange);
- market-related characteristics (active and sizable market, low volatility, flight to quality).

High quality liquid asset

Table: Stock of HQLA

Level	Description	Haircut
Level 1 assets		
	Coins and bank notes Sovereign, central bank, PSE, and MDB assets qualifying for 0% risk weighting Central bank reserves Domestic sovereign or central bank debt for non-0% risk weighting	0%
Level 2 assets (maximum of 40% of HQLA)		
Level 2A assets		
	Sovereign, central bank, PSE and MDB assets qualifying for 20% risk weighting Corporate debt securities rated AA− or higher Covered bonds rated AA− or higher	15%
Level 2B assets (maximum of 15% of HQLA)		
	RMBS rated AA or higher	25%
	Corporate debt securities rated between A+ and BBB−	50%
	Common equity shares	50%

High quality liquid asset

Level 2 assets are subject to two caps. Let x_{HQLA}, x_1 and x_2 be the value of HQLA, level 1 assets and level 2 assets. We have:

$$\begin{array}{rcl} x_{\mathrm{HQLA}} & = & x_1 + x_2 \\ & \text{s.t.} & \left\{ \begin{array}{l} x_2 = x_{2A} + x_{2B} \\ x_{2A} \leq 0.40 \cdot x_{\mathrm{HQLA}} \\ x_{2B} \leq 0.15 \cdot x_{\mathrm{HQLA}} \end{array} \right. \end{array}$$

We deduce that one trivial solution is:

$$\left\{ \begin{array}{l} x_{\mathrm{HQLA}}^{\star} = \min\left(\dfrac{5}{3}x_1, x_1 + x_2\right) \\ x_1^{\star} = x_1 \\ x_2^{\star} = x_{\mathrm{HQLA}}^{\star} - x_1^{\star} \\ x_{2A}^{\star} = \min\left(x_2^{\star}, x_{2A}\right) \\ x_{2B}^{\star} = x_2^{\star} - x_{2A}^{\star} \end{array} \right.$$

High quality liquid asset

Example

We consider the following assets:

1. Coins and bank notes = \$200 mn
2. Central bank reserves = \$100 mn
3. 20% risk-weighted sovereign debt securities = \$200 mn
4. AA corporate debt securities = \$300 mn
5. Qualifying RMBS = \$200 mn
6. BB+ corporate debt securities = \$500 mn

High quality liquid asset

Table: Solution of the exercise

	Assets	Gross Value	Haircut	Net Value	Capped Value
Level 1 assets	(1) + (2)	300	0%	300	300
Level 2 assets		1 200		825	200
2A	(3) + (4)	500	15%	425	200
2B	(5) + (6)	700		400	0
	(5)	200	25%	150	0
	(6)	500	50%	250	0
Total		1 500		1 125	500

$\Rightarrow$ The stock of HQLA is equal to \$500 mn

Total net cash outflows

The value of total net cash outflows is defined as follows:

$$\text{Total net cash outflows} = \text{Total expected cash outflows} - \min\left(\begin{array}{c}\text{Total expected cash inflows,} \\ \text{75\% of total expected cash outflows}\end{array}\right)$$

Total net cash outflows

Table: Cash outflows of the LCR

Liabilities	Description	Rate
Retail deposits		
Demand and term deposits (less than 30 days)		
	Stable deposits covered by deposit insurance	3%
	Stable deposits	5%
	Less stable deposits	10%
Term deposits (with residual maturity greater than 30 days)		0%
Unsecured wholesale funding		
Demand and term deposits (less than 30 days) provided by small business customers		
	Stable deposits	5%
	Less stable deposits	10%
Deposits generated by clearing, custody and cash management		25%
	Portion covered by deposit insurance	5%
Cooperative banks in an institutional network		25%
Corporates, sovereigns, central banks, PSEs and MDBs		40%
	Portion covered by deposit insurance	20%

Total net cash outflows

Table: Cash outflows of the LCR

Liabilities Description	Rate
Secured funding transactions	
With a central bank counterparty	0%
Backed by level 1 assets	0%
Backed by level 2A assets	15%
Backed by non-level 1 or non-level 2A assets with domestic sovereigns, PSEs or MDBs as a counterparty	25%
Backed by level 2B RMBS assets	25%
Backed by other level 2B assets	50%
All other secured funding transactions	100%
Additional requirements	
Margin/collateral calls	$\geq$ 20%
ABCP, SIVs, conduits, SPVs, etc.	100%
Net derivative cash outflows	100%
Other credit/liquidity facilities	$\geq$ 5%

Total net cash outflows

Table: Cash inflows of the LCR

Receivables Description	Rate
Maturing secured lending transactions	
Backed by level 1 assets	0%
Backed by level 2A assets	15%
Backed by level 2B RMBS assets	25%
Backed by other level 2B assets	50%
Backed by non-HQLAs	100%
Other cash inflows	
Credit/liquidity facilities provided to the bank	0%
Inflows to be received from retail counterparties	50%
Inflows to be received from non-financial wholesale counterparties	50%
Inflows to be received from financial institutions and central banks	100%
Net derivative receivables	100%

Total net cash outflows

Example

The bank has \$500 mn of HQLA. Its main liabilities are:

1. Retail stable deposit = \$17.8 bn (\$15 bn have a government guarantee)
2. Retail term deposit (with a maturity of 6 months) = \$5 bn
3. Stable deposit provided by small business customers = \$1 bn
4. deposit of corporates = \$200 mn

In the next thirty days, the bank also expects to receive \$100 mn of loan repayments, and \$10 mn due to a maturing derivative

Total net cash outflows

- We calculate the expected cash outflows for the next thirty days:

$$\begin{aligned}\text{Cash outflows} &= 3\% \times 15\,000 + 5\% \times 2\,800 + 0\% \times 5\,000 + \\ &\quad 5\% \times 1\,000 + 40\% \times 200 \\ &= \$720 \text{ mn}\end{aligned}$$

- We estimate the cash inflows expected by the bank for the next month:

$$\text{Cash inflows} = 50\% \times 100 + 100\% \times 10 = \$60 \text{ mn}$$

- We deduce that the liquidity coverage ratio of the bank is equal to:

$$\mathrm{LCR} = \frac{500}{720 - 60} = 75.76\%$$

Net stable funding ratio

It is defined as the amount of available stable funding (ASF) relative to the amount of required stable funding (RSF):

$$\mathrm{NSFR} = \frac{\text{Available amount of stable funding}}{\text{Required amount of stable funding}} \geq 100\%$$

- The available amount of stable funding (ASF) corresponds to the regulatory capital plus some other liabilities
- The required amount of stable funding (RSF) is the sum of weighted assets and off-balance sheet exposures

Leverage ratio

- It is defined as the capital measure divided by the exposure measure
- This ratio must be below 3%
- The capital measure corresponds to the tier 1 capital
- The exposure measure is composed of four main exposures:
    1. On-balance sheet exposures
    2. Derivative exposures
    3. Securities financing transaction (SFT)
    4. Exposures and off-balance sheet items

References

Basel Committee on Banking Supervision (2013)
Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools, January 2013.

Basel Committee on Banking Supervision (2014)
Basel III: The Net Stable Funding Ratio, October 2014.

Basel Committee on Banking Supervision (2017)
Basel III: Finalising Post-crisis Reforms, December 2017.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 6.

Course 2023-2024 in Financial Risk Management
Lecture 7. Asset Liability Management Risk

Thierry Roncalli*

*Amundi Asset Management[16]

*University of Paris-Saclay

September 2023

[16]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- ALM risk $\Rightarrow$ banking book
- Not only a risk management issue, but also concerns commercial choices and business models
- Several ALM risks: liquidity risk, interest rate risk, embedded option risk, currency risk

The ALM function is located in the finance department, not in the risk management department

Balance sheet

Table: A simplified balance sheet

Assets	Liabilities
Cash	Due to central banks
Loans and leases	Deposits
Mortgages	Deposit accounts
Consumer credit	Savings
Credit cards	Term deposits
Interbank loans	Interbank funding
Investment securities	Short-term debt
Sovereign bonds	Subordinated debt
Corporate bonds	Reserves
Other assets	Equity capital

Income statement

Table: A simplified income statement

	Interest income
−	Interest expenses
=	Net interest income
+	Non-interest income
=	Gross income
−	Operating expenses
=	Net income
−	Provisions
=	Earnings before tax
−	Income tax
=	Profit after tax

Accounting standards

Four main systems:

1. US GAAP
2. Japanese combined system
3. Chinese accounting standards
4. International Financial Reporting Standards (IFRS)

$\Rightarrow$ IFRS is implemented in European Union, Australia, Middle East, Russia, South Africa, etc.

Accounting standards

Before 2018

IAS 39

- financial assets at fair value through profit and loss (FVTPL)
- available-for-sale financial assets (AFS)
- loans and receivables (L&R);
- held-to-maturity investments (HTM)

After 2018

IFRS 9

- amortized cost (AC)
- fair value (FV)

ALM committee (ALCO)

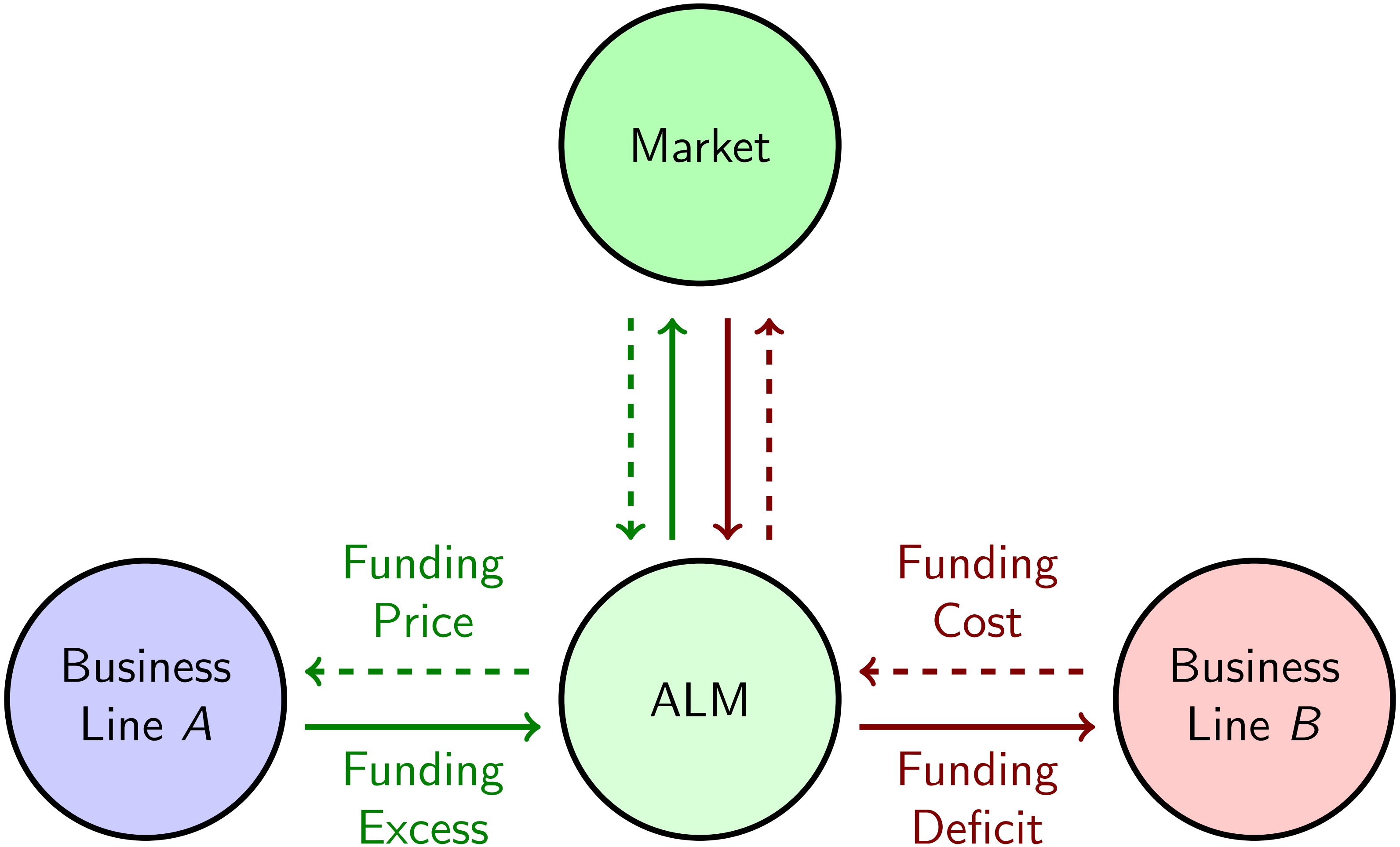


Figure: Internal and external funding transfer

Liquidity gap

- $A(t)$ is the value of assets at time t
- $L(t)$ is the value of liabilities at time t
- Funding ratio

$$\mathcal{FR}(t) = \frac{A(t)}{L(t)}$$

- Funding gap

$$\mathcal{FG}(t) = A(t) - L(t)$$

- Liquidity ratio

$$\mathcal{LR}(t) = \frac{L(t)}{A(t)}$$

- Liquidity gap

$$\mathcal{LG}(t) = L(t) - A(t)$$

Liquidity gap

Example

The assets $A(t)$ are composed of loans that are linearly amortized in a monthly basis during the next year. The amount is equal to 120. The liabilities $L(t)$ are composed of three short-term in fine debt instruments, and the capital. The corresponding debt notional is respectively equal to 65, 10 and 5 whereas the associated remaining maturity is equal to two, seven and twelve months. The amount of capital is stable for the next twelve months and is equal to 40

Liquidity gap

Table: Computation of the liquidity gap

Period	0	1	2	3	4	5	6	7	8	9	10	11	12
Loans	120	110	100	90	80	70	60	50	40	30	20	10	0
Assets	120	110	100	90	80	70	60	50	40	30	20	10	0
Debt #1	65	65	65										
Debt #2	10	10	10	10	10	10	10	10					
Debt #3	5	5	5	5	5	5	5	5	5	5	5	5	5
Debt (total)	80	80	80	15	15	15	15	15	5	5	5	5	5
Equity	40	40	40	40	40	40	40	40	40	40	40	40	40
Liabilities	120	120	120	55	55	55	55	55	45	45	45	45	45
$\mathcal{LG}(t)$	0	10	20	-35	-25	-15	-5	5	5	15	25	35	45

Liquidity gap

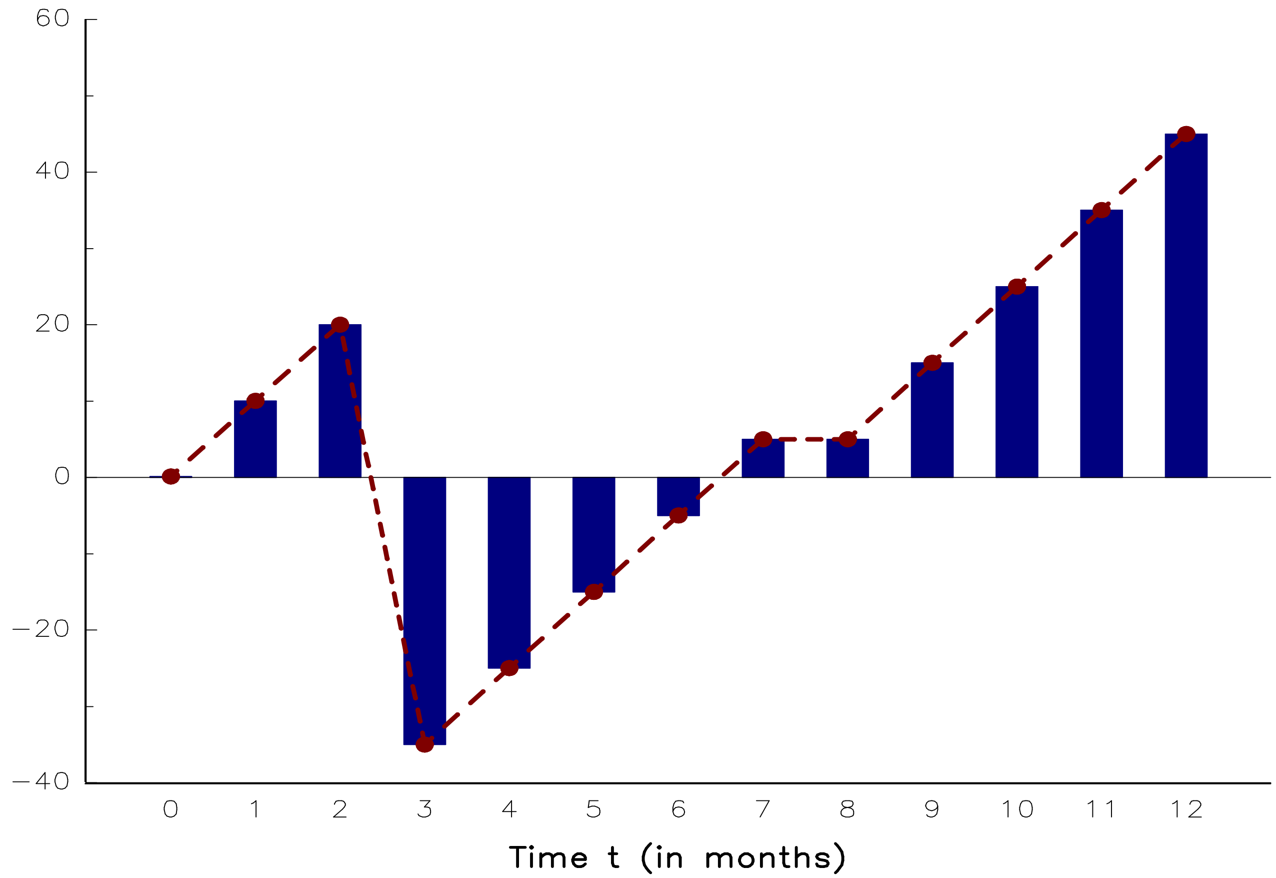


Figure: An example of liquidity gap

Asset and liability amortization

General principles of debt amortization

- The annuity amount $A(t)$ at time t is composed of the interest payment $I(t)$ and the principal payment $P(t)$:

$$A(t) = I(t) + P(t)$$

- The interest payment at time t is equal to the interest rate $i(t)$ times the outstanding principal balance $N(t-1)$:

$$I(t) = i(t) N(t-1)$$

- The outstanding principal balance $N(t)$ equal to

$$N(t) = N(t-1) - P(t)$$

- The outstanding principal balance $N(t)$ is equal to the present value $C(t)$ of forward annuity amounts: $N(t) = C(t)$

Asset and liability amortization

- Constant amortization debt (or linear amortization of the capital): $P(t)$ is constant over time (HFRM, page 379):

$$\begin{aligned} P(t) &= \frac{1}{n} N_0 \\ A(t) &= I(t) + P(t) = \left(\frac{1}{n} + i \left(1 - \frac{t-1}{n} \right) \right) N_0 \end{aligned}$$

- Constant payment debt: the annuity amount $A(t)$ is constant

$$\begin{aligned} A(t) &= A = \frac{i}{1-(1+i)^{-n}} N_0 \\ I(t) &= \left(1 - \frac{1}{(1+i)^{n-t+1}} \right) A \end{aligned}$$

- Bullet repayment debt: the notional is fully repaid at the time of maturity

$$I(t) = iN_0 \text{ and } P(t) = \mathbb{1}\{t = n\} \cdot N_0$$

Asset and liability amortization

Example

We consider a 10-year mortgage, whose notional is equal to \$100. The annual interest rate i is equal to 5%, and we assume annual principal payments

Asset and liability amortization

Table: Repayment schedule of the constant amortization mortgage

t	$C(t-1)$	$A(t)$	$I(t)$	$P(t)$	$Q(t)$	$N(t)$
1	100.00	15.00	5.00	10.00	10.00	90.00
2	90.00	14.50	4.50	10.00	20.00	80.00
3	80.00	14.00	4.00	10.00	30.00	70.00
4	70.00	13.50	3.50	10.00	40.00	60.00
5	60.00	13.00	3.00	10.00	50.00	50.00
6	50.00	12.50	2.50	10.00	60.00	40.00
7	40.00	12.00	2.00	10.00	70.00	30.00
8	30.00	11.50	1.50	10.00	80.00	20.00
9	20.00	11.00	1.00	10.00	90.00	10.00
10	10.00	10.50	0.50	10.00	100.00	0.00

Asset and liability amortization

Table: Repayment schedule of the constant payment mortgage

t	$C(t-1)$	$A(t)$	$I(t)$	$P(t)$	$Q(t)$	$N(t)$
1	100.00	12.95	5.00	7.95	7.95	92.05
2	92.05	12.95	4.60	8.35	16.30	83.70
3	83.70	12.95	4.19	8.77	25.06	74.94
4	74.94	12.95	3.75	9.20	34.27	65.73
5	65.73	12.95	3.29	9.66	43.93	56.07
6	56.07	12.95	2.80	10.15	54.08	45.92
7	45.92	12.95	2.30	10.65	64.73	35.27
8	35.27	12.95	1.76	11.19	75.92	24.08
9	24.08	12.95	1.20	11.75	87.67	12.33
10	12.33	12.95	0.62	12.33	100.00	0.00

Asset and liability amortization

Table: Repayment schedule of the bullet repayment mortgage

t	$C(t-1)$	$A(t)$	$I(t)$	$P(t)$	$Q(t)$	$N(t)$
1	100.00	5.00	5.00	0.00	0.00	100.00
2	100.00	5.00	5.00	0.00	0.00	100.00
3	100.00	5.00	5.00	0.00	0.00	100.00
4	100.00	5.00	5.00	0.00	0.00	100.00
5	100.00	5.00	5.00	0.00	0.00	100.00
6	100.00	5.00	5.00	0.00	0.00	100.00
7	100.00	5.00	5.00	0.00	0.00	100.00
8	100.00	5.00	5.00	0.00	0.00	100.00
9	100.00	5.00	5.00	0.00	0.00	100.00
10	100.00	105.00	5.00	100.00	100.00	0.00

Asset and liability amortization

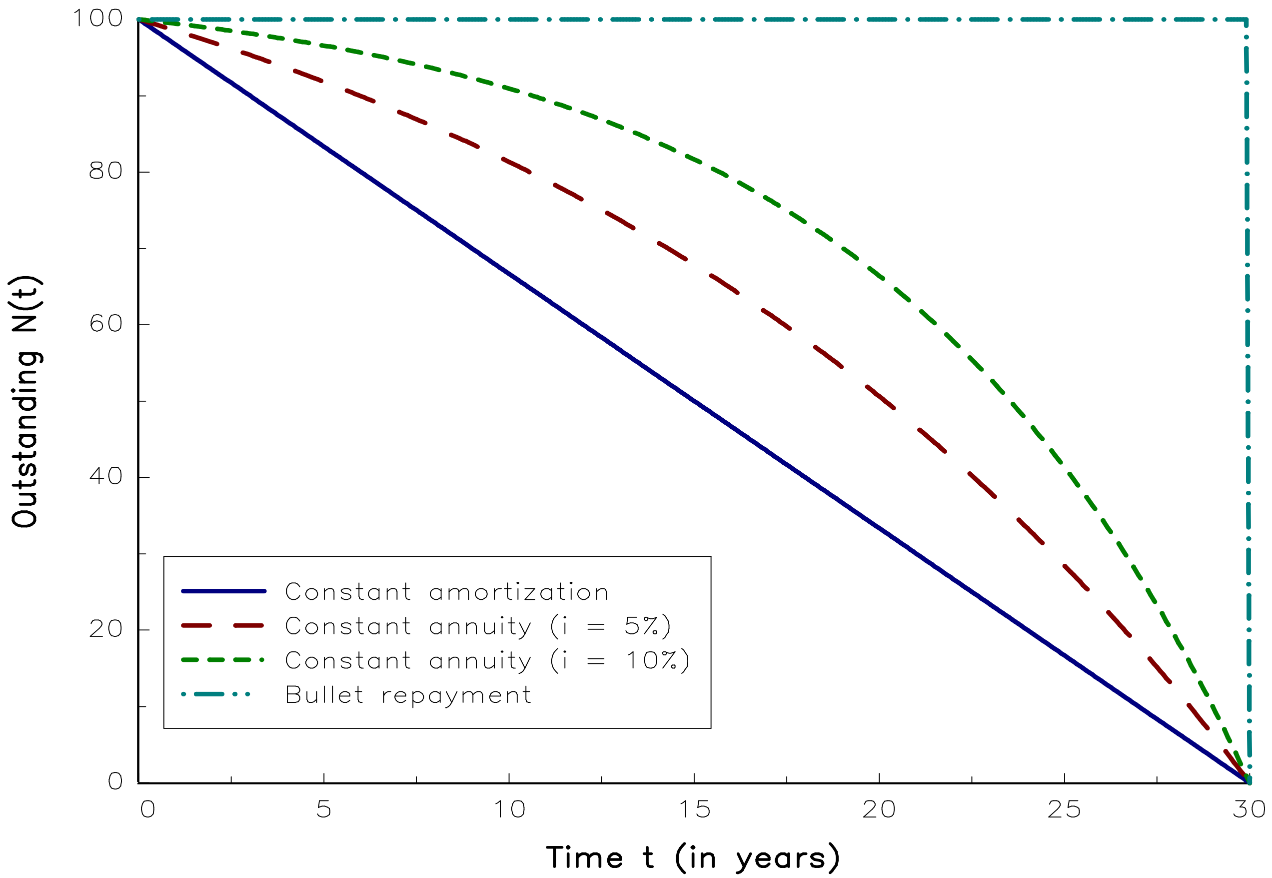


Figure: Amortization schedule of the 30-year mortgage (monthly payments)

Asset and liability amortization

Example

We consider the following balance sheet:

Assets				Liabilities			
Items	Notional	Rate	Mat.	Items	Notional	Rate	Mat.
Loan #1	100	5%	10	Debt #1	120	5%	10
Loan #2	50	8%	16	Debt #2	80	3%	5
Loan #3	40	3%	8	Debt #3	70	4%	10
Loan #4	110	2%	7	Capital #4	30		

All the debt instruments are subject to monthly principal payments

Mixed schedule = constant principal (loan #3 and debt #2), constant annuity (loan #1, loan #2 and debt #1) and bullet repayment (loan #4 and debt #2)

Asset and liability amortization

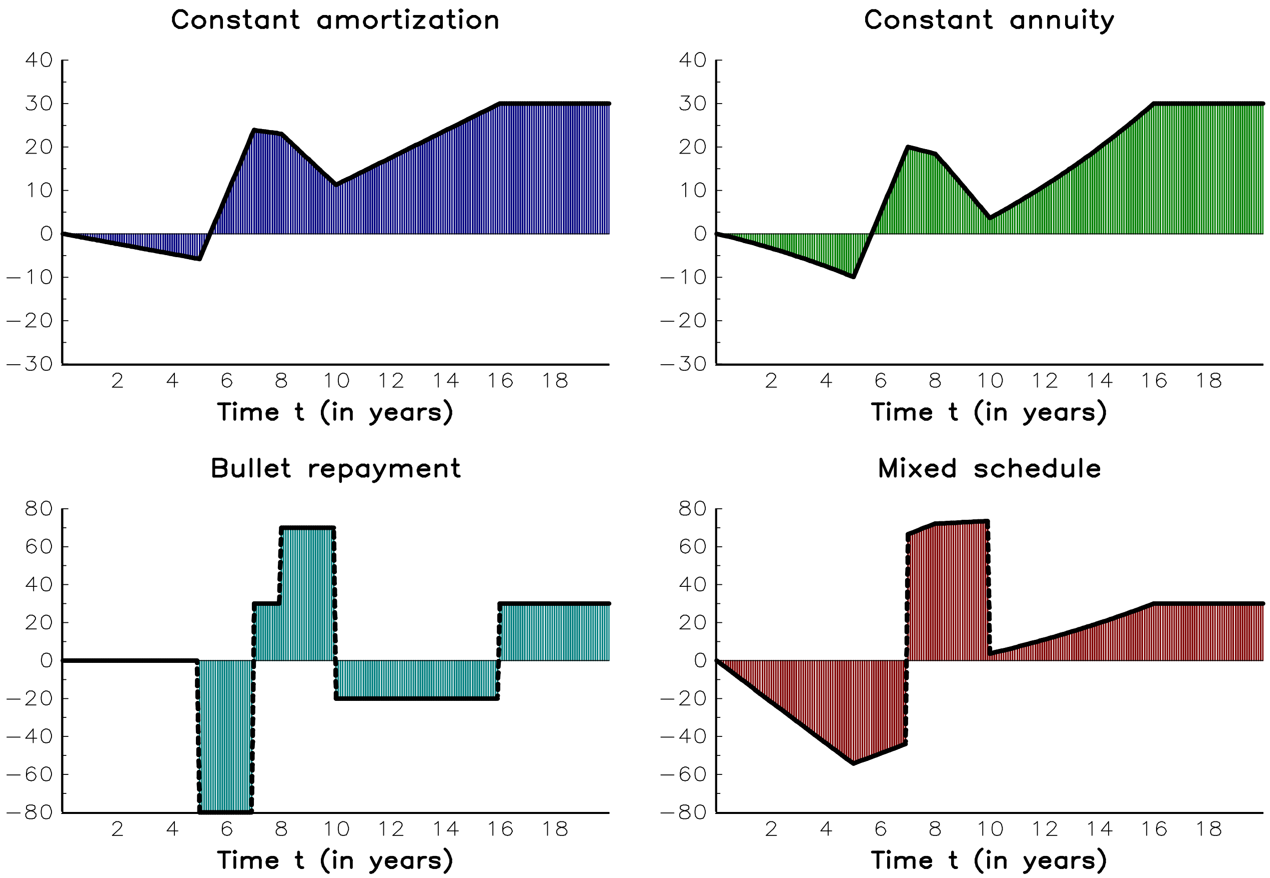


Figure: Impact of the amortization schedule on the liquidity gap

Asset and liability amortization

Table: Computation of the liquidity gap (mixed schedule, **first twelve months**)

t	Assets					Liabilities					$\mathcal{LG}_t$
	#1	#2	#3	#4	A_t	#1	#2	#3	#4	L_t	
1	99.4	49.9	39.6	110	298.8	119.2	78.7	70	30	297.9	−0.92
2	98.7	49.7	39.2	110	297.6	118.5	77.3	70	30	295.8	−1.83
3	98.1	49.6	38.8	110	296.4	117.7	76.0	70	30	293.7	−2.75
4	97.4	49.5	38.3	110	295.2	116.9	74.7	70	30	291.6	−3.66
5	96.8	49.3	37.9	110	294.0	116.1	73.3	70	30	289.4	−4.58
6	96.1	49.2	37.5	110	292.8	115.3	72.0	70	30	287.3	−5.49
7	95.4	49.1	37.1	110	291.6	114.5	70.7	70	30	285.2	−6.41
8	94.8	48.9	36.7	110	290.4	113.7	69.3	70	30	283.1	−7.32
9	94.1	48.8	36.3	110	289.2	112.9	68.0	70	30	280.9	−8.24
10	93.4	48.7	35.8	110	287.9	112.1	66.7	70	30	278.8	−9.15
11	92.8	48.5	35.4	110	286.7	111.3	65.3	70	30	276.7	−10.06
12	92.1	48.4	35.0	110	285.5	110.5	64.0	70	30	274.5	−10.97

Asset and liability amortization

Table: Computation of the liquidity gap (mixed schedule, **annual schedule**)

t	Assets					Liabilities					$\mathcal{LG}_t$
	#1	#2	#3	#4	A_t	#1	#2	#3	#4	L_t	
0	100.0	50.0	40.0	110	300.0	120.0	80.0	70	30	300.0	0.00
1	92.1	48.4	35.0	110	285.5	110.5	64.0	70	30	274.5	−10.97
2	83.8	46.7	30.0	110	270.4	100.5	48.0	70	30	248.5	−21.90
3	75.0	44.8	25.0	110	254.8	90.1	32.0	70	30	222.1	−32.76
4	65.9	42.7	20.0	110	238.6	79.0	16.0	70	30	195.0	−43.55
5	56.2	40.5	15.0	110	221.7	67.4		70	30	167.4	−54.27
6	46.1	38.1	10.0	110	204.2	55.3		70	30	155.3	−48.91
7	35.4	35.5	5.0		75.9	42.5		70	30	142.5	66.56
8	24.2	32.7			56.9	29.0		70	30	129.0	72.12
9	12.4	29.7			42.1	14.9		70	30	114.9	72.81
10		26.4			26.4				30	30.0	3.62
11		22.8			22.8				30	30.0	7.19
12		18.9			18.9				30	30.0	11.06
13		14.8			14.8				30	30.0	15.24
14		10.2			10.2				30	30.0	19.77
15		5.3			5.3				30	30.0	24.68
16					0.0				30	30.0	30.00

Impact of prepayment

We have:

$$N^c(t) = N(t) \cdot \mathbb{1}\{\boldsymbol{\tau} > t\}$$

where:

- $N^c(t)$ and $N(t)$ are the outstanding principal balances with and without prepayment
- $\boldsymbol{\tau}$ is the prepayment time of the debt instrument

We deduce that:

$$\mathbb{E}[N^c(t)] = \mathbf{S}(t) \cdot N(t)$$

where $\mathbf{S}(t) = \mathbb{E}[\mathbb{1}\{\boldsymbol{\tau} > t\}]$ is the survival function of $\boldsymbol{\tau}$

Remark

If $\boldsymbol{\tau} \sim \mathcal{E}(\lambda)$ where λ is the prepayment intensity, we obtain:

$$\mathbb{E}[N^c(t)] = e^{-\lambda t} \cdot N(t) \leq N(t)$$

Impact of prepayment

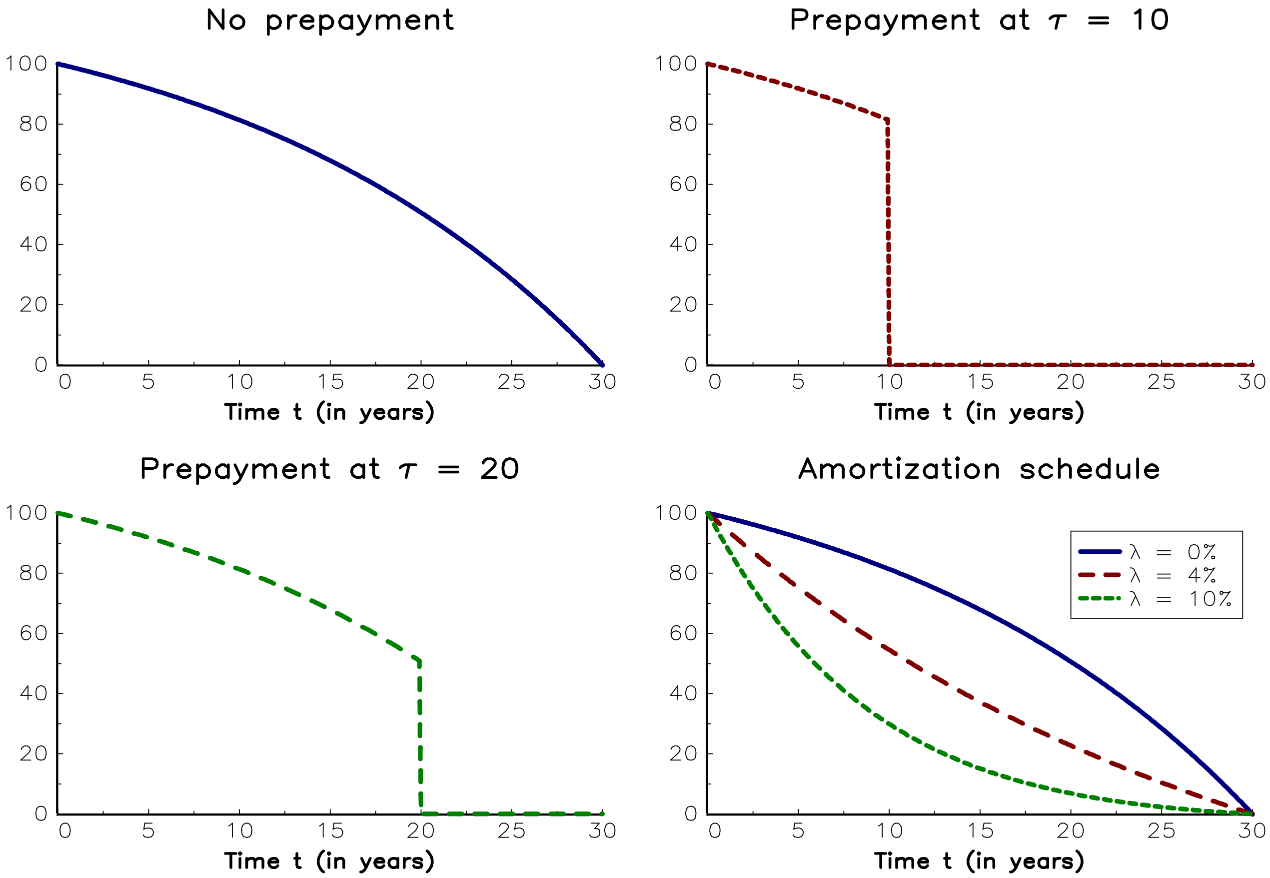


Figure: Conventional amortization schedule with prepayment risk

Impact of new production

Accounting identity

$$N(t) = N(t-1) - \mathrm{AM}(t) + \mathrm{NP}(t)$$

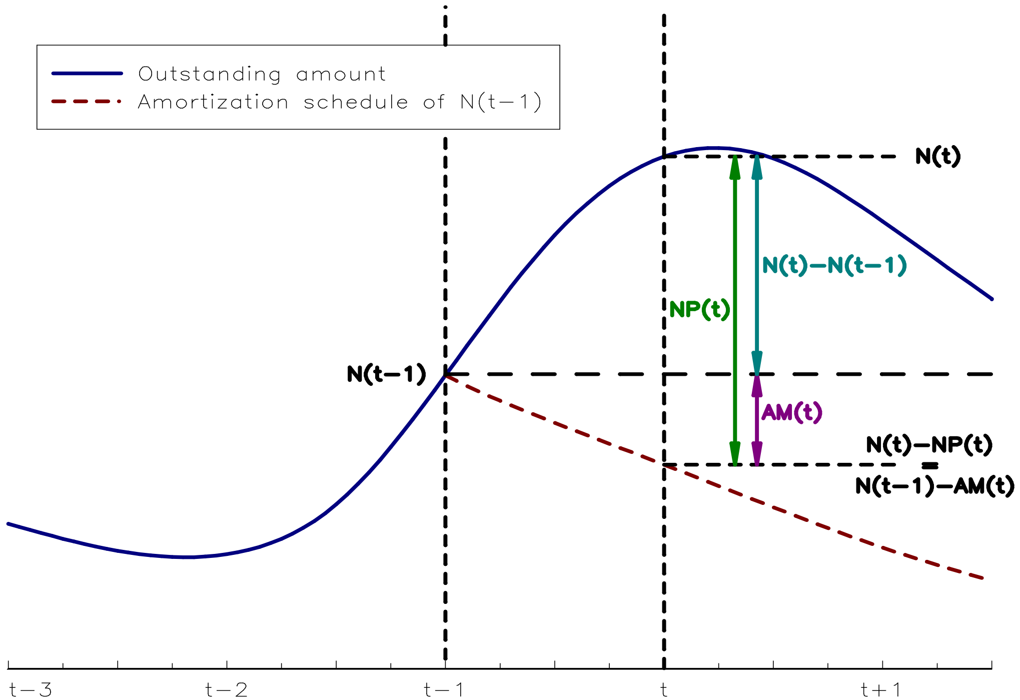


Figure: Impact of the new production on the outstanding amount

Dynamic analysis

- Run-off balance sheet
  A balance sheet where existing non-trading book positions amortize and are not replaced by any new business.
- Constant balance sheet
  A balance sheet in which the total size and composition are maintained by replacing maturing or repricing cash flows with new cash flows that have identical features.
- Dynamic balance sheet
  A balance sheet incorporating future business expectations, adjusted for the relevant scenario in a consistent manner.

Dynamic analysis
Notations

- $\text{NP}(t)$: New production at time t
- $\text{NP}(t,u)$: New production at time t that is present in the balance sheet at time $u \geq t$
- $\mathbf{S}(t,u)$: Survival function of the new production
- $f(t,u)$ is the density function associated to the survival function $\mathbf{S}(t,u)$
- $N(t,u)$: Non-amortized outstanding amount at time t that is present in the balance sheet at time $u \geq t$
- $\mathbf{S}^{\star}(t,u)$: Survival function of the outstanding amount

Dynamic analysis

Stock-flow analysis

- The amortization function $\mathbf{S}(t,u)$ is defined by:

$$\mathrm{NP}(t,u) = \mathrm{NP}(t) \cdot \mathbf{S}(t,u)$$

It measures the proportion of \$1 entering in the balance sheet at time t that remains present at time $u \geq t$:

$$N(t) = \int_{-\infty}^{t} \mathrm{NP}(s)\, \mathbf{S}(s,t)\, \mathrm{d}s$$

- The amortization function $\mathbf{S}^{\star}(t,u)$ is defined by:

$$N(t,u) = N(t) \cdot \mathbf{S}^{\star}(t,u)$$

It measures the proportion of \$1 of outstanding amount at time t that remains present at time $u \geq t$

$$\mathbf{S}^{\star}(t,u) = \frac{\int_{-\infty}^{t} \mathrm{NP}(s)\, \mathbf{S}(s,u)\, \mathrm{d}s}{\int_{-\infty}^{t} \mathrm{NP}(s)\, \mathbf{S}(s,t)\, \mathrm{d}s}$$

Dynamic analysis

Dynamics of the outstanding amount

We have:

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = -\int_{-\infty}^{t} \mathrm{NP}(s)\, f(s,t)\, \mathrm{d}s + \mathrm{NP}(t)$$

where $f(t,u) = -\partial_u \mathbf{S}(t,u)$ is the density function of the amortization

Dynamic analysis

Estimation of the dynamic liquidity gap

The dynamic liquidity gap at time t for a future date $u \geq t$ is given by:

$$\begin{aligned} \mathcal{LG}(t,u) &= \sum_{k \in \mathcal{L}iabilities} \left(N_k(t,u) + \int_t^u \mathrm{NP}_k(s)\, \mathbf{S}_k(s,u)\, \mathrm{d}s \right) - \\ & \quad \sum_{k \in \mathcal{A}ssets} \left(N_k(t,u) + \int_t^u \mathrm{NP}_k(s)\, \mathbf{S}_k(s,u)\, \mathrm{d}s \right) \end{aligned}$$

In the case of the run-off balance sheet, we obtain:

$$\mathcal{LG}(t,u) = \sum_{k \in \mathcal{L}iabilities} N_k(t,u) - \sum_{k \in \mathcal{A}ssets} N_k(t,u)$$

Dynamic analysis
Liquidity duration

The liquidity duration is the weighted average life (WAL) of the principal repayments:

$$\mathcal{D}(t) = \int_t^{\infty} (u - t) f(t, u) \, \mathrm{d}u$$

where $f(t, u)$ is the density function associated to the survival function $\mathbf{S}(t, u)$

Remark

All the previous formulas can be obtained in the discrete-time analysis (HFRM, Section 7.1.2.3, pages 385-392)

Dynamic analysis
Mathematical formulas

Table: Survival function and liquidity duration of some amortization schemes (HFRM, Exercise 7.4.3, page 450; HFRM-CB, Section 7.4.3, pages 126-128)

Amortization	$\mathbf{S}(t,u)$	$\mathcal{D}(t)$
Bullet	$\mathbb{1}\{t \leq u < t+m\}$	m
Constant	$\mathbb{1}\{t \leq u < t+m\} \cdot \left(1 - \frac{u-t}{m}\right)$	$\frac{m}{2}$
Exponential	$e^{-\lambda(u-t)}$	$\frac{1}{\lambda}$
Amortization	$\mathbf{S}^{\star}(t,u)$	$\mathcal{D}^{\star}(t)$
Bullet	$\mathbb{1}\{t \leq u < t+m\} \cdot \left(1 - \frac{u-t}{m}\right)$	$\frac{m}{2}$
Constant	$\mathbb{1}\{t \leq u < t+m\} \cdot \left(1 - \frac{u-t}{m}\right)^2$	$\frac{m}{3}$
Exponential	$e^{-\lambda(u-t)}$	$\frac{1}{\lambda}$
Amortization	$\mathrm{d}N(t)$	
Bullet	$\mathrm{d}N(t) = (\mathrm{NP}(t) - \mathrm{NP}(t-m))\,\mathrm{d}t$	
Constant	$\mathrm{d}N(t) = \left(\mathrm{NP}(t) - \frac{1}{m}\int_{t-m}^{t} \mathrm{NP}(s)\,\mathrm{d}s\right)\mathrm{d}t$	
Exponential	$\mathrm{d}N(t) = (\mathrm{NP}(t) - \lambda N(t))\,\mathrm{d}t$	

Dynamic analysis
Illustration

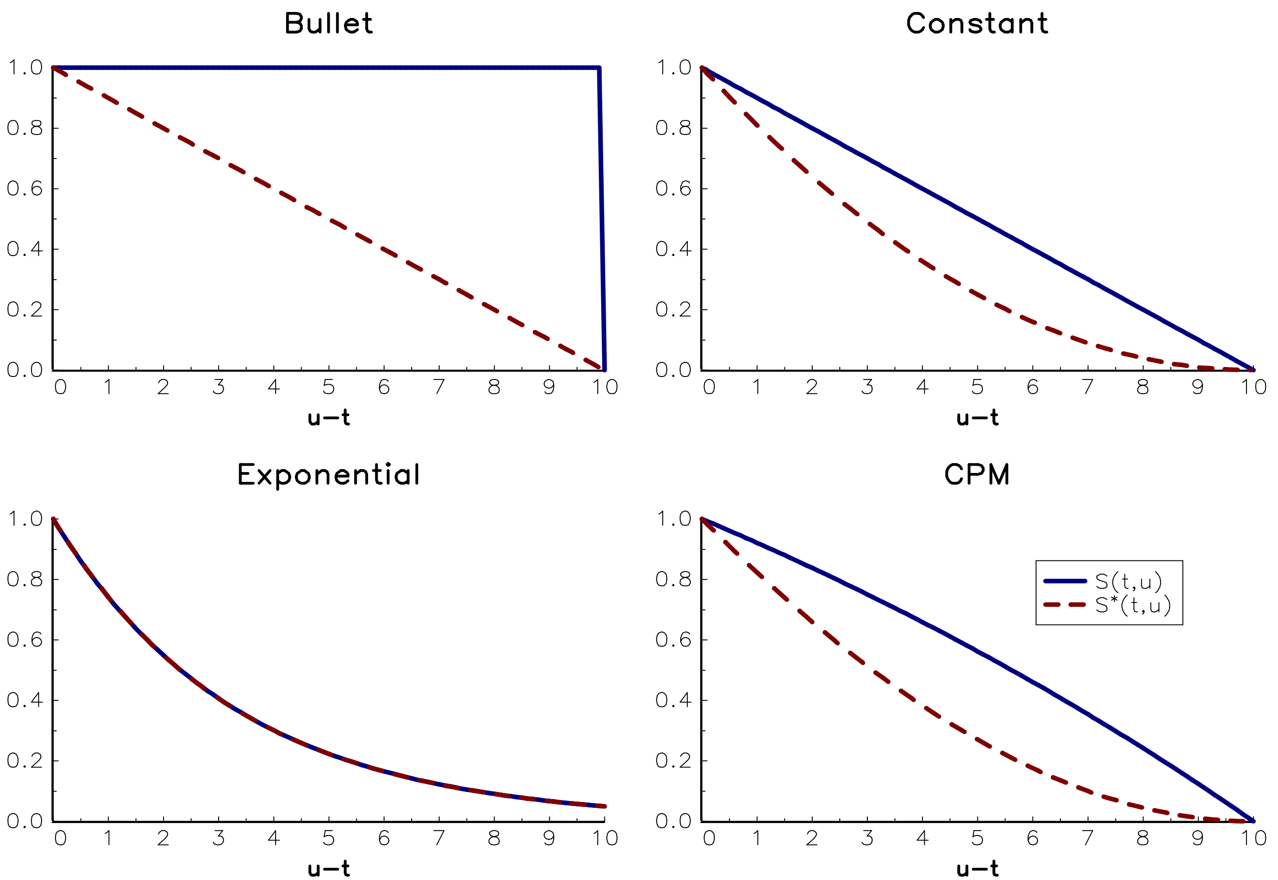


Figure: Amortization functions $\mathbf{S}(t,u)$ and $\mathbf{S}^{\star}(t,u)$

Definition

"IRRBB refers to the current or prospective risk to the bank' capital and earnings arising from adverse movements in interest rates that affect the bank's banking book positions. When interest rates change, the present value and timing of future cash flows change. This in turn changes the underlying value of a bank's assets, liabilities and off-balance sheet items and hence its economic value. Changes in interest rates also affect a bank's earnings by altering interest rate-sensitive income and expenses, affecting its net interest income" (BCBS, 2016)

1. Economic value (EV, EVE): changes in the net present value of the balance sheet
2. Earnings-based risk measures (EaR, NII): changes in the expected future profitability of the bank

Categories of IRR

3 main sources of interest rate risk

- Gap risk: mismatch risk arising from the term structure of banking book instruments
  - Repricing risk
  - Yield curve risk
- Basis risk: mismatch risk arising from different interest rate indices
  - Correlation risk of interest rate indices with the same maturity
- Option risk: option derivative positions and embedded options
  - Automatic option risk (caps, floors, swaptions and other interest rate derivatives)
  - Behavioral option risk
    - Prepayment risk
    - Early redemption risk (or withdrawal risk)
    - Non-maturity deposit (NMD)

Economic value (EV)

The economic value of a series of cash flows $\mathrm{CF} = \{\mathrm{CF}(t_k), t_k \geq t\}$ is the present value of these cash flows:

$$\mathrm{EV} = \mathbb{E}\left[\sum_{t_k \geq t} \mathrm{CF}(t_k) \cdot e^{-\int_t^{t_k} r(s)\,\mathrm{d}s}\right] = \sum_{t_k \geq t} \mathrm{CF}(t_k) \cdot B(t, t_k)$$

where $B(t, t_k)$ is the discount factor for the maturity date t_k

Application to the banking book

- We slot all notional repricing cash flows of assets and liabilities into a set of time buckets
- We calculate the net cash flows:

$$\mathrm{CF}(t_k) = \mathrm{CF}_A(t_k) - \mathrm{CF}_L(t_k)$$

where $\mathrm{CF}_A(t_k)$ and $\mathrm{CF}_L(t_k)$ are the cash flows of assets and liabilities for the time bucket t_k

- The economic value is given by:

$$\begin{aligned} \mathrm{EV} &= \sum_{t_k \geq t} \mathrm{CF}(t_k) \cdot B(t, t_k) \\ &= \sum_{t_k \geq t} \mathrm{CF}_A(t_k) \cdot B(t, t_k) - \sum_{t_k \geq t} \mathrm{CF}_L(t_k) \cdot B(t, t_k) \\ &= \mathrm{EV}_A - \mathrm{EV}_L \end{aligned}$$

Stress testing of the economic value

- We note EV_s the economic value corresponding to the stress scenario s
- EV_0 is the base scenario and corresponds to the current term structure of interest rates
- We have:

$$\begin{aligned}\Delta \mathrm{EV}_s &= \mathrm{EV}_0 - \mathrm{EV}_s \\ &= \sum_{t_k \geq t} \mathrm{CF}_0(t_k) \cdot B_0(t, t_k) - \sum_{t_k \geq t} \mathrm{CF}_s(t_k) \cdot B_s(t, t_k)\end{aligned}$$

Economic value of equity (EVE)

The economic value of equity (EVE or EV_E) is a specific form of EV where equity is excluded from the cash flows

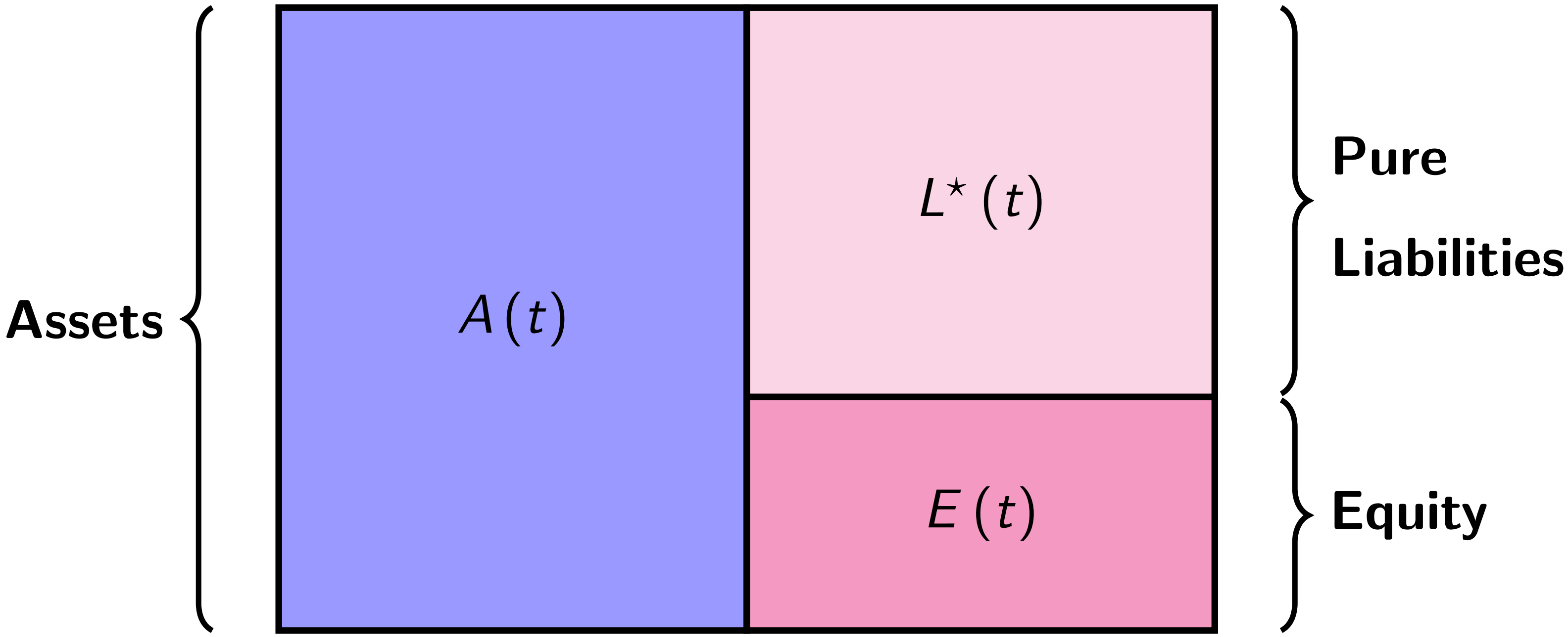


Figure: Relationship between $A(t)$, $L^{\star}(t)$ and $E(t)$

We have:

$$A(t) = L(t) = L^{\star}(t) + E(t)$$

Economic value of equity

Since the value of the capital is equal to $E(t) = A(t) - L^{\star}(t)$, we have:

$$\mathrm{EVE} = \mathrm{EV}_A - \mathrm{EV}_{L^{\star}}$$

and:

$$\Delta\,\mathrm{EVE}_s = \mathrm{EVE}_0 - \mathrm{EVE}_s$$

Remark

The economic value of equity is then equal to:

$$\mathrm{EVE} = \sum_{t_k \geq t} \mathrm{CF}_A(t_k) \cdot B(t, t_k) - \sum_{t_k \geq t} \mathrm{CF}_{L^{\star}}(t_k) \cdot B(t, t_k)$$

Net interest income (NII)

- The net interest income is the difference between the interest payments on assets and the interest payments of liabilities
- We have:

$$\Delta\,\mathrm{NII}_s = \mathrm{NII}_0 - \mathrm{NII}_s$$

- $\Delta\,\mathrm{NII}_s > 0$ indicates a loss if the stress scenario s occurs

Basel III

The risk measures are equal to the maximum of losses by considering the different scenarios:

$$\mathcal{R}(\mathrm{EVE}) = \max_s (\Delta\,\mathrm{EVE}_s, 0)$$

and:

$$\mathcal{R}(\mathrm{NII}) = \max_s (\Delta\,\mathrm{NII}_s, 0)$$

IRRBB

- No minimum capital requirements $\mathcal{K}$
- $\mathcal{R}(\mathrm{EVE}) \leq 15\% \times \mathrm{Tier}\,1$
- Pillar 2

Interest rate risk principles

- 9 IRR principles for banks (management, risk appetite, model governance process, capital adequacy policy, etc.)
- 3 IRR principles for supervisors (data collection, challenging the model assumptions, identification of outlier banks)

Some examples:

- To compute $\Delta \mathrm{EVE}$, banks must consider a run-off balance sheet assumption
- To compute $\Delta \mathrm{NII}$, banks must use a constant or dynamic balance sheet and a rolling 12-month period
- Banks must use:
  - Internal (historical and hypothetical) interest rate scenarios
  - 6 external interest rate scenarios

The standardized approach

5 steps for measuring the bank's IRRBB

1. The first step consists in allocating the interest rate sensitivities of the banking book to three categories
   1. amenable to standardization
   2. less amenable to standardization
   3. not amenable to standardization
2. Then, the bank must slot cash flows into 19 predefined time buckets: overnight (O/N), O/N–1M, ..., 10Y–15Y, 15Y–20Y, 20Y+
3. The bank determines $\Delta \mathrm{EVE}_{s,c}$ for each shock s and each currency c
4. The bank calculates the total measure for automatic interest rate option risk $\mathrm{KAO}_{s,c}$
5. The bank calculates the EVE for each shock s:

$$\mathcal{R}(\mathrm{EVE}_s) = \max\left(\sum_c \left(\Delta \mathrm{EVE}_{s,c} + \mathrm{KAO}_{s,c}\right)^+ ; 0\right)$$

The standardized EVE risk measure is equal to:

$$\mathcal{R}(\mathrm{EVE}) = \max_s \mathcal{R}(\mathrm{EVE}_s)$$

The standardized approach

The supervisory interest rate shock scenarios

Three shock sizes:

Shock size		USD/CAD/SEK	EUR/HKD	GBP	JPY	EM
$\mathbb{S}_0$	(parallel)	200	200	250	100	400
$\mathbb{S}_1$	(short)	300	250	300	100	500
$\mathbb{S}_2$	(long)	150	100	150	100	300

Given $\mathbb{S}_0$, $\mathbb{S}_1$ and $\mathbb{S}_2$, we calculate the following generic shocks for a given maturity t:

	Parallel shock $\Delta R^{(\mathrm{parallel})}(t)$	Short rates shock $\Delta R^{(\mathrm{short})}(t)$	Long rates shock $\Delta R^{(\mathrm{long})}(t)$
Up	$+\mathbb{S}_0$	$+\mathbb{S}_1 \cdot e^{-t/\tau}$	$+\mathbb{S}_2 \cdot \left(1 - e^{-t/\tau}\right)$
Down	$-\mathbb{S}_0$	$-\mathbb{S}_1 \cdot e^{-t/\tau}$	$-\mathbb{S}_2 \cdot \left(1 - e^{-t/\tau}\right)$

where τ is equal to four years

The standardized approach

The supervisory interest rate shock scenarios

The six standardized interest rate shock scenarios are defined as follows:

1. Parallel shock up: $\Delta R^{(\text{parallel})}(t) = +\mathbb{S}_0$
2. Parallel shock down: $\Delta R^{(\text{parallel})}(t) = -\mathbb{S}_0$
3. Steepener shock (short rates down and long rates up):
$$\Delta R^{(\text{steepnener})}(t) = 0.90 \cdot \left|\Delta R^{(\text{long})}(t)\right| - 0.65 \cdot \left|\Delta R^{(\text{short})}(t)\right|$$
4. Flattener shock (short rates up and long rates down):
$$\Delta R^{(\text{flattener})}(t) = 0.80 \cdot \left|\Delta R^{(\text{short})}(t)\right| - 0.60 \cdot \left|\Delta R^{(\text{long})}(t)\right|$$
5. Short rates shock up:
$$\Delta R^{(\text{short})}(t) = +\mathbb{S}_1 \cdot e^{-t/\tau}$$
6. Short rates shock down:
$$\Delta R^{(\text{short})}(t) = -\mathbb{S}_1 \cdot e^{-t/\tau}$$

The standardized approach

The supervisory interest rate shock scenarios

Example

We assume that $\mathbb{S}_0 = 100$ bps, $\mathbb{S}_1 = 150$ bps and $\mathbb{S}_2 = 200$ bps. We would like to calculate the standardized shocks for the one-year maturity

- The parallel shock up is equal to $+100$ bps, while the parallel shock down is equal to -100 bps
- For the short rates shock, we obtain:

$$\Delta R^{(\text{short})}(t) = 150 \times e^{-1/4} = 116.82 \text{ bps}$$

for the up scenario and -116.82 bps for the down scenario
- Since we have $\left|\Delta R^{(\text{short})}(t)\right| = 116.82$ and $\left|\Delta R^{(\text{long})}(t)\right| = 44.24$, the steepener shock is equal to:

$$\Delta R^{(\text{steepnener})}(t) = 0.90 \times 44.24 - 0.65 \times 116.82 = -36.12 \text{ bps}$$

For the flattener shock, we have:

$$\Delta R^{(\text{flattener})}(t) = 0.80 \times 116.82 - 0.60 \times 44.24 = 66.91 \text{ bps}$$

The standardized approach

The supervisory interest rate shock scenarios

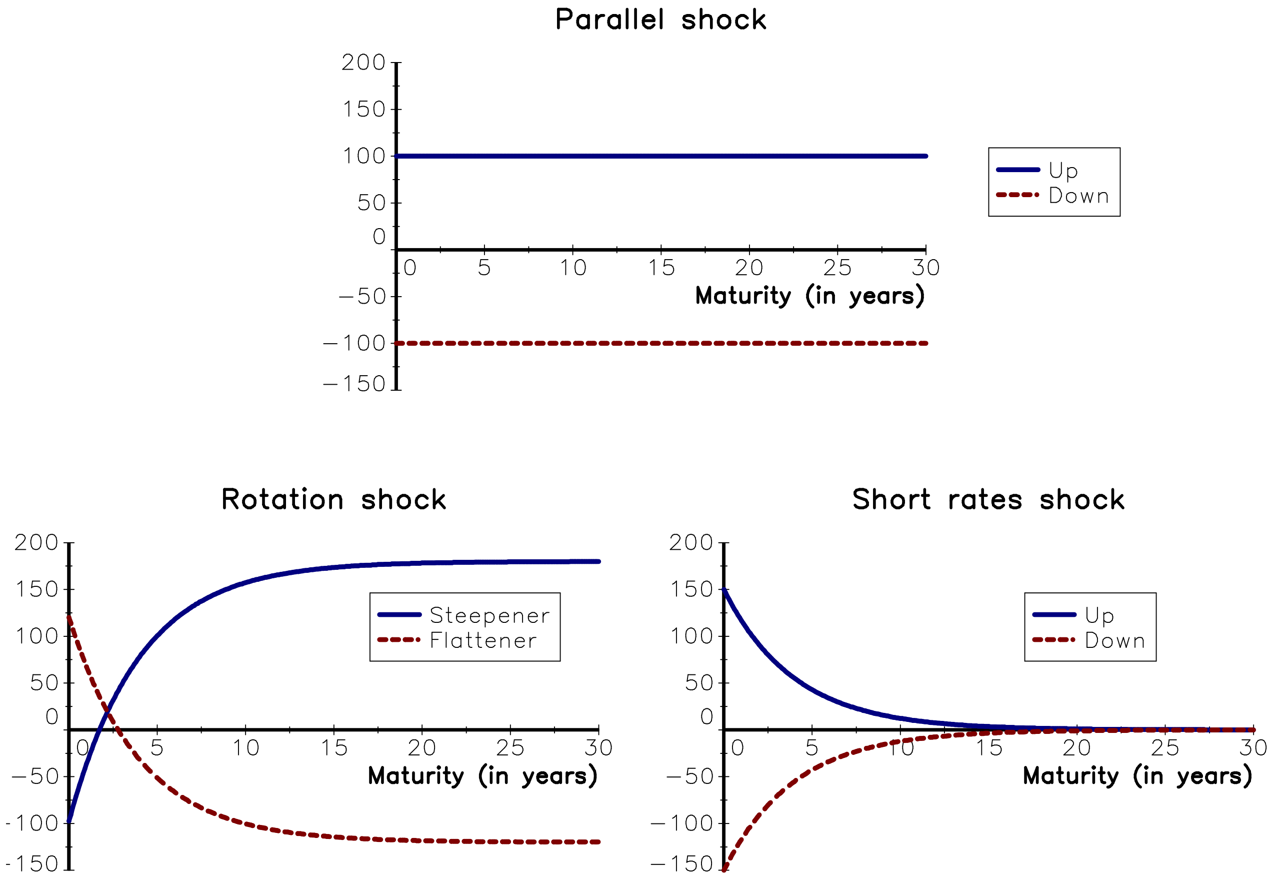


Figure: Interest rate shocks (in bps) with $(\mathbb{S}_0 = 100, \mathbb{S}_1 = 150, \mathbb{S}_2 = 200)$

The standardized approach

The supervisory interest rate shock scenarios

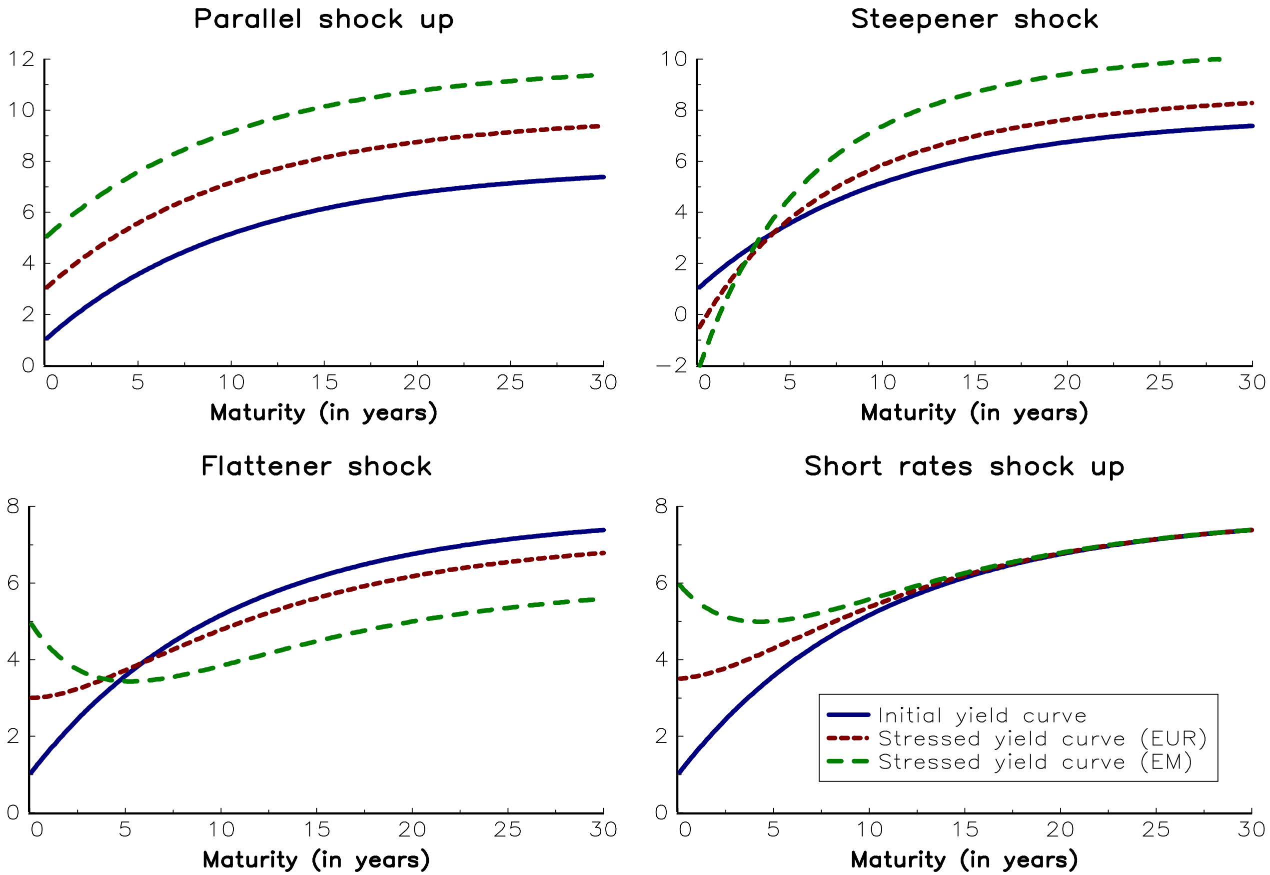


Figure: Stressed yield curve (in %)

The standardized approach

Treatment of NMDs

- Retail transactional (RT)
- Retail non-transactional (RNT)
- Wholesale (W)

- Difference between stable and non-stable part of each category
- The stable part of NMDs must be split between:
  - Core deposits
    - Maximum proportion: 90% for RT, 70% for RNT and 50% for W
    - Maximum maturity: 5Y for RT, 4.5Y for RNT and 4Y for W
  - Non-core deposits (overnight maturity)

The standardized approach

Prepayment risk

The bank estimates the baseline conditional prepayment rate (CPR) CPR_0 and calculates the stressed conditional prepayment rate as follows:

$$\mathrm{CPR}_s = \min\left(1, \gamma_s \cdot \mathrm{CPR}_0\right)$$

where γ_s is the multiplier for the scenario s and

- $\gamma_s = 0.8$ for the scenarios 1, 3 and 5 (parallel up, steepener and short rates up)
- $\gamma_s = 1.2$ for the scenarios 2, 4 and 6 (parallel down, flattener and short rates down)

The standardized approach

Early redemption

The term deposit redemption ratio (TDRR) is stressed as follows:

$$\mathrm{TDRR}_s = \min\left(1, \gamma_s \cdot \mathrm{TDRR}_0\right)$$

where γ_s is the multiplier for the scenario s and:

- $\gamma_s = 1.2$ for the scenarios 1, 4 and 5 (parallel up, flattener and short rates up)
- $\gamma_s = 0.8$ for the scenarios 2, 3 and 6 (parallel down, steepener and short rates down)

The standardized approach

Automatic interest rate options

The computation of the automatic interest rate option risk KAO_s is given by:

$$\mathrm{KAO}_s = \sum_{i \in \mathcal{S}} \Delta \mathrm{FVAO}_{s,i} - \sum_{i \in \mathcal{B}} \Delta \mathrm{FVAO}_{s,i}$$

where:

- $i \in \mathcal{S}$ (resp. $i \in \mathcal{B}$) denotes an automatic interest rate option which is sold (resp. bought) by the bank
- $\mathrm{FVAO}_{s,i}$ (resp. $\mathrm{FVAO}_{0,i}$) is the fair value of the automatic option i given the stressed (resp. current) yield curve and a relative increase in the implied volatility of 25% (resp. the current implied volatility)
- $\Delta \mathrm{FVAO}_{s,i}$ is the change in the fair value of the option:

$$\Delta \mathrm{FVAO}_{s,i} = \mathrm{FVAO}_{s,i} - \mathrm{FVAO}_{0,i}$$

The standardized approach

Example

We consider a USD-denominated balance sheet. The assets are composed of loans with the following cash flow slotting:

Instruments	Loans	Loans	Loans
Maturity	1Y	5Y	13Y
Cash flows	200	700	200

The liabilities are composed of retail deposit accounts, term deposits, debt and tier-one equity capital:

Instruments	Non-core deposits	Term deposits	Core deposits	Debt ST	Debt LT	Equity capital
Maturity	O/N	7M	3Y	4Y	8Y	
Cash flows	100	50	450	100	100	200

The standardized approach
Example

Table: Economic value of the assets

Bucket	t_k	$\mathrm{CF}_0(t_k)$	$R_0(t_k)$	$\mathrm{EV}_0(t_k)$
6	0.875	200	1.55%	197.31
11	4.50	700	3.37%	601.53
17	12.50	100	5.71%	48.98
EV_0				847.82

Table: Economic value of the pure liabilities

Bucket	t_k	$\mathrm{CF}_0(t_k)$	$R_0(t_k)$	$\mathrm{EV}_0(t_k)$
1	0.0028	100	1.00%	100.00
5	0.625	50	1.39%	49.57
9	2.50	450	2.44%	423.35
10	3.50	100	2.93%	90.26
14	7.50	100	4.46%	71.56
EV_0				734.73

The standardized approach

Example

Table: Stressed economic value of equity

Bucket	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$
			Assets			
$R_s(t_6)$	3.55%	−0.45%	0.24%	3.30%	3.96%	−0.87%
$R_s(t_{11})$	5.37%	1.37%	3.65%	3.54%	4.34%	2.40%
$R_s(t_{17})$	7.71%	3.71%	6.92%	4.96%	5.84%	5.58%
$\mathrm{EV}_s(t_6)$	193.89	200.80	199.57	194.31	193.20	201.52
$\mathrm{EV}_s(t_{11})$	549.76	658.18	594.03	596.91	575.74	628.48
$\mathrm{EV}_s(t_{17})$	38.15	62.89	42.13	53.83	48.18	49.79
EV_s	781.79	921.87	835.74	845.05	817.11	879.79
			Pure liabilities			
$R_s(t_1)$	3.00%	−1.00%	−0.95%	3.40%	4.00%	−2.00%
$R_s(t_5)$	3.39%	−0.61%	−0.08%	3.32%	3.96%	−1.17%
$R_s(t_9)$	4.44%	0.44%	2.03%	3.31%	4.05%	0.84%
$R_s(t_{10})$	4.93%	0.93%	2.90%	3.40%	4.18%	1.68%
$R_s(t_{14})$	6.46%	2.46%	5.31%	4.07%	4.92%	4.00%
$\mathrm{EV}_s(t_1)$	99.99	100.00	100.00	99.99	99.99	100.01
$\mathrm{EV}_s(t_5)$	48.95	50.19	50.02	48.97	48.78	50.37
$\mathrm{EV}_s(t_9)$	402.70	445.05	427.77	414.27	406.69	440.69
$\mathrm{EV}_s(t_{10})$	84.16	96.80	90.34	88.77	86.39	94.30
$\mathrm{EV}_s(t_{14})$	61.59	83.14	67.17	73.70	69.13	74.07
EV_s	697.39	775.18	735.31	725.71	710.98	759.43
			Equity			
EVE_s	84.41	146.68	100.43	119.34	106.13	120.37
$\Delta\,\mathrm{EVE}_s$	28.69	−33.58	12.67	−6.24	6.97	−7.27

The standardized approach

Example

- The current economic value of equity is equal to:

$$\mathrm{EVE}_0 = 847.82 - 734.73 = 113.09$$

- In the case of the first stress scenario, we have:

$$\mathrm{EVE}_1 = 781.79 - 697.39 = 84.41$$

  and:

$$\Delta\,\mathrm{EVE}_1 = 113.10 - 84.41 = 28.69$$

- EVE decreases for scenarios 1, 3 and 5
- The EVE risk measure is equal to:

$$\mathcal{R}\,(\mathrm{EVE}) = \max_s \left(\Delta\,\mathrm{EVE}_s, 0\right) = 28.69$$

  It represents 14.3% of the equity:

$$\frac{28.69}{200} = 14.3\%$$

The materiality test is not satisfied

Currency risk

- Currency hedging $\Rightarrow$ also the equity capital?
- Dollar funding
- Multi-currency balance sheet

Credit spread risk

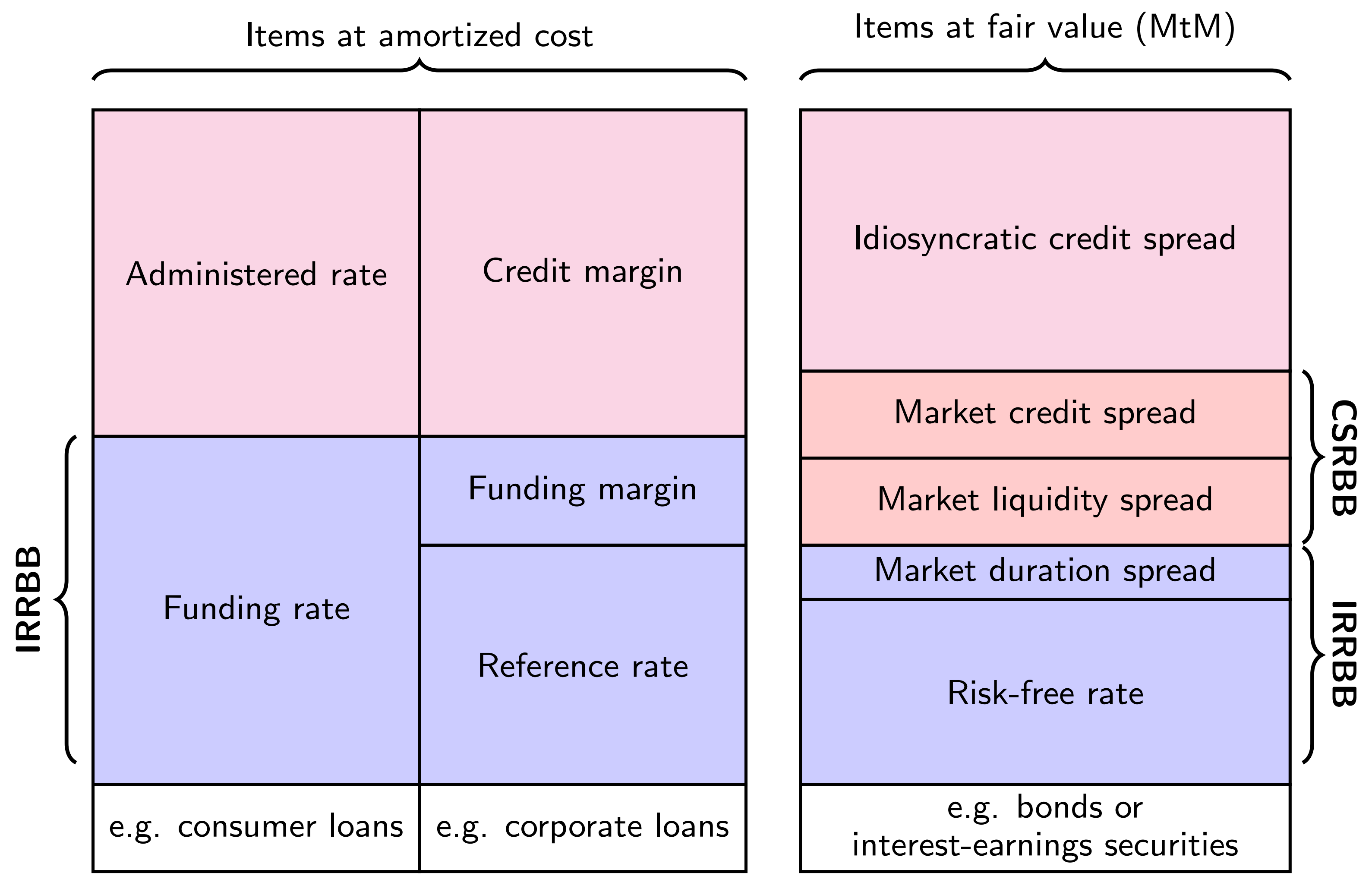


Figure: Components of interest rates

Macaulay duration

The Macaulay duration $\mathcal{D}$ is the weighted average of the cash flow maturities:

$$\mathcal{D} = \sum_{t_k \geq t} w(t, t_k) \cdot (t_k - t)$$

We have:

$$\frac{\partial V}{\partial y} = -\frac{\mathcal{D}}{1+y} \cdot V = -\mathfrak{D} \cdot V$$

where $\mathfrak{D}$ is the modified duration

Application to a portfolio

The market value of the portfolio is composed of m cash flow streams: $V = \sum_{j=1}^{m} V_j$ while the duration of the portfolio is the average of individual durations: $\mathcal{D} = \sum_{j=1}^{m} w_j \cdot \mathcal{D}_j$ where $w_j = \frac{V_j}{V}$

Duration gap risk

Since $E(t) = A(t) - L^{\star}(t)$ and $\mathrm{EV}_E = \mathrm{EV}_A - \mathrm{EV}_{L^{\star}}$, the duration of equity is equal to:

$$\mathcal{D}_E = \frac{\mathrm{EV}_A}{\mathrm{EV}_A - \mathrm{EV}_{L^{\star}}} \cdot \mathcal{D}_A - \frac{\mathrm{EV}_{L^{\star}}}{\mathrm{EV}_A - \mathrm{EV}_{L^{\star}}} \cdot \mathcal{D}_{L^{\star}} = \frac{\mathrm{EV}_A}{\mathrm{EV}_A - \mathrm{EV}_{L^{\star}}} \cdot \mathcal{D}_{\mathcal{G}ap}$$

where the duration gap (also called DGAP) is equal to

$$\mathcal{D}_{\mathcal{G}ap} = \mathcal{D}_A - \frac{\mathrm{EV}_{L^{\star}}}{\mathrm{EV}_A} \cdot \mathcal{D}_{L^{\star}}$$

Duration gap risk

Another expression of the equity duration is:

$$\mathcal{D}_E = \frac{\mathrm{EV}_A}{\mathrm{EV}_E} \cdot \mathcal{D}_{\mathcal{G}ap} = \mathcal{L}_{A/E} \cdot \mathcal{D}_{\mathcal{G}ap}$$

where $\mathcal{L}_{A/E}$ is the leverage ratio

Relationship between EVE and duration gap

$$\begin{aligned} \Delta \,\mathrm{EVE} &= \Delta \,\mathrm{EV}_E \\ &\approx -\mathcal{D}_E \cdot \mathrm{EV}_E \cdot \frac{\Delta y}{1+y} \\ &\approx -\mathcal{D}_{\mathcal{G}ap} \cdot \mathrm{EV}_A \cdot \frac{\Delta y}{1+y} \end{aligned}$$

Illustration

We consider the following balance sheet:

Assets	V_j	$\mathcal{D}_j$	Liabilities	V_j	$\mathcal{D}_j$
Cash	5	0.0	Deposits	40	3.2
Loans	40	1.5	CDs	20	0.8
Mortgages	40	6.0	Debt	30	1.7
Securities	15	3.8	Equity capital	10	
Total	100		Total	100	

We have $\mathrm{EV}_A = 100$, $\mathrm{EV}_{L^\star} = 90$, $\mathrm{EV}_E = 10$ and:

$$\mathcal{L}_{A/E} = \frac{\mathrm{EV}_A}{\mathrm{EV}_E} = \frac{100}{10} = 10$$

The duration values are equal to:

$$\begin{aligned}
\mathcal{D}_A &= \frac{5}{100} \times 0 + \frac{40}{100} \times 1.5 + \frac{40}{100} \times 6.0 + \frac{15}{100} \times 3.8 = 3.57 \text{ years} \\
\mathcal{D}_{L^\star} &= \frac{40}{90} \times 3.2 + \frac{20}{90} \times 0.8 + \frac{30}{90} \times 1.7 = 2.17 \text{ years}
\end{aligned}$$

Illustration

We deduce that:

$$\mathcal{D}_{\mathcal{G}ap} = 3.57 - \frac{90}{100} \times 2.17 = 1.62 \text{ years}$$

If we assume that the current yield to maturity is equal to 3%, we obtain:

Δy	-2%	-1%	$+1\%$	$+2\%$
$\Delta\,\mathrm{EVE}$	3.15	1.57	−1.57	−3.15
$\dfrac{\Delta\,\mathrm{EVE}}{\mathrm{EVE}}$	31.46%	15.73%	−15.73%	−31.46%

Immunization of the balance sheet

We must have:

$$\Delta \,\mathrm{EVE} = 0 \Leftrightarrow \mathcal{D}_{\mathcal{G}ap} = 0 \Leftrightarrow \mathcal{D}_A - \frac{\mathrm{EV}_{L^\star}}{\mathrm{EV}_A} \cdot \mathcal{D}_{L^\star} = 0$$

Table: Bank balance sheet after immunization of the duration gap

Assets	V_j	$\mathcal{D}_j$	Liabilities	V_j	$\mathcal{D}_j$
Cash	5	0.0	Deposits	40	3.2
Loans	40	1.5	CDs	20	0.8
Mortgages	40	6.0	Debt	10.48	1.7
Securities	15	3.8	Zero-coupon bond	19.52	10.0
			Equity capital	10	0.0
Total	100		Total	100	

Income gap analysis

- If interest rates change, this induces a gap (or repricing) risk because the bank will have to reinvest assets and refinance liabilities at a different interest rate level in the future
- The gap is defined as the difference between rate sensitive assets (RSA) and rate sensitive liabilities (RSL):

$$\mathrm{GAP}\left(t,u\right)=\mathrm{RSA}\left(t,u\right)-\mathrm{RSL}\left(t,u\right)$$

  where t is the current date and u is the time horizon of the gap
- We can show that:

$$\Delta\,\mathrm{NII}\left(t,u\right)\approx\mathrm{GAP}\left(t,u\right)\cdot\Delta R$$

  where ΔR is the parallel shock of interest rates

Net interest income

The net interest income of the bank is the difference between interest rate revenues of its assets and interest rate expenses of its liabilities:

$$\mathrm{NII}(t,u) = \sum_{i \in \mathcal{A}ssets} N_i(t,u) \cdot R_i(t,u) - \sum_{j \in \mathcal{L}iabilities} N_j(t,u) \cdot R_j(t,u)$$

where $\mathrm{NII}(t,u)$ is the net interest income at time t for the maturity date u

⇒ Mathematical formulation (HFRM, pages 412-418)

Modeling customer rates

Client rates $\neq$ market rates

Several issues:

- Correlation
- Next repricing date (known or unknown?)
- Sensitivity of the customer rate with respect to the market rate

Hedging strategies

Using a forward rate agreement, we can show that:

$$\mathrm{NII}_{\mathcal{H}}(t,u) - \mathrm{NII}(t,u) = \mathrm{GAP}(t,u) \cdot \rho(t,u) \cdot (f(t,u) - r(u))$$

We can draw several conclusions:

- When the interest rate gap is closed, the bank does not need to hedge the net interest income
- When the correlation $\rho(t,u)$ between the customer rate and the market rate is equal to one, the notional of the hedge is exactly equal to the interest rate gap (it is lower in the general case)
- If the bank hedges the net interest income and if the gap is positive, a decrease of interest rates is not favorable

Hedging the interest rate gap depends on the expectations of the bank $\Longrightarrow$ partial hedging and macro hedging

Hedging instruments

- Interest rate swaps (IRS)
- Forward rate agreements (FRA)
- Swaptions

Funds transfer pricing

All liquidity and interest rate risks are transferred to the ALM unit:

- Business units can then lend or borrow funding at a given internal price
- This price is called the funds transfer price (FTP) or the internal transfer rate
- The FTP charges interests to the business unit for client loans, whereas the FTP compensates the business unit for raising deposits

Net interest margin

- The net interest margin (NIM) is equal to:

$$\begin{aligned}\mathrm{NIM}(t,u) &= \frac{\sum_{i \in \mathcal{A}ssets} N_i(t,u) \cdot R_i(t,u) - \sum_{j \in \mathcal{L}iabilities} N_j(t,u) \cdot R_j(t,u)}{\sum_{i \in \mathcal{A}ssets} N_i(t,u)} \\ &= \frac{\mathrm{RA}(t,u) \cdot R_{\mathrm{RA}}(t,u) - \mathrm{RL}(t,u) \cdot R_{\mathrm{RL}}(t,u)}{\mathrm{RA}(t,u)}\end{aligned}$$

  where R_{RA} and R_{RL} represent the weighted average interest rate of interest earning assets and interest bearing liabilities

- The net interest spread (NIS) is equal to:

$$\begin{aligned}\mathrm{NIS}(t,u) &= \frac{\sum_{i \in \mathcal{A}ssets} N_i(t,u) \cdot R_i(t,u)}{\sum_{i \in \mathcal{A}ssets} N_i(t,u)} - \frac{\sum_{j \in \mathcal{L}iabilities} N_j(t,u) \cdot R_j(t,u)}{\sum_{j \in \mathcal{L}iabilities} N_j(t,u)} \\ &= R_{\mathrm{RA}}(t,u) - R_{\mathrm{RL}}(t,u)\end{aligned}$$

- NIM is the profitability ratio of the assets whereas NIS is the interest rate spread captured by the bank

Net interest margin

Example

We consider the following interest earning and bearing items:

Assets	$N_i(t,u)$	$R_i(t,u)$	Liabilities	$N_j(t,u)$	$R_j(t,u)$
Loans	100	5%	Deposits	100	0.5%
Mortgages	100	4%	Debts	60	2.5%

Net interest margin

- The interest income is equal to $100 \times 5\% + 100 \times 4\% = 9$ whereas the interest expense is equal to $100 \times 0.5\% + 60 \times 2.5\% = 2$. We deduce that the net interest income is equal to:

$$\mathrm{NII}(t,u) = 9 - 2 = 7$$

- We have:

$$R_{\mathrm{RA}}(t,u) = \frac{100 \times 5\% + 100 \times 4\%}{100 + 100} = 4.5\%$$

and:

$$R_{\mathrm{RL}}(t,u) = \frac{100 \times 0.5\% + 60 \times 2.5\%}{100 + 60} = 1.25\%$$

Since $\mathrm{RA}(t,u) = 200$ and $\mathrm{RL}(t,u) = 160$, we deduce that:

$$\mathrm{NIM}(t,u) = \frac{200 \times 4.5\% - 160 \times 1.25\%}{200} = \frac{7}{200} = 3.5\%$$

and:

$$\mathrm{NIS}(t,u) = 4.5\% - 1.25\% = 3.25\%$$

Net interest margin

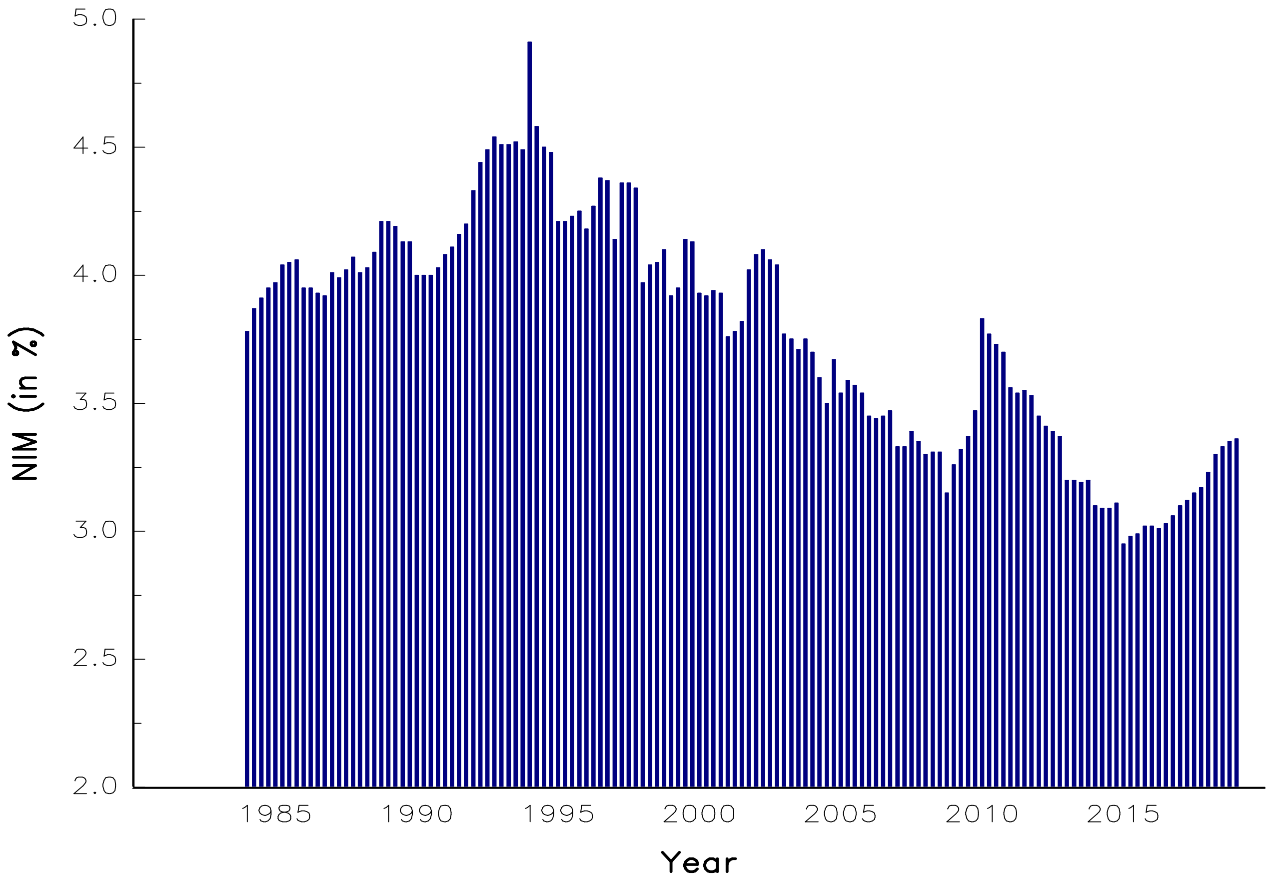


Figure: Evolution of the net interest margin in the US

Commercial margin

- The commercial margin rate is the spread between the client rate $R_i(t,u)$ of the asset i and the corresponding market rate $r(t,u)$:

$$m_i(t,u) = R_i(t,u) - r(t,u)$$

- For the liability j, we have:

$$m_j(t,u) = r(t,u) - R_j(t,u)$$

- In the case where we can perfectly match the asset i with the liability j, the commercial margin rate is the net interest spread:

$$m(t,u) = m_i(t,u) + m_j(t,u) = R_i(t,u) - R_j(t,u)$$

Commercial margin

Introducing a funds transfer pricing system is equivalent to interpose the ALM unit between the business unit and the market

- For assets, we have:

$$m_i(t,u) = \underbrace{\left(R_i(t,u) - \mathrm{FTP}_i(t,u)\right)}_{m_i^{(c)}(t,u)} + \underbrace{\left(\mathrm{FTP}_i(t,u) - r(t,u)\right)}_{m_i^{(t)}(t,u)}$$

  where:
  - $m_i^{(c)}(t,u)$ is the commercial margin rate of the business unit
  - $m_i^{(t)}(t,u)$ is the transformation margin rate of the ALM unit
- For liabilities, we have:

$$m_j(t,u) = \underbrace{\left(\mathrm{FTP}_j(t,u) - R_j(t,u)\right)}_{m_j^{(c)}(t,u)} + \underbrace{\left(r(t,u) - \mathrm{FTP}_j(t,u)\right)}_{m_j^{(t)}(t,u)}$$

Commercial margin

Commercial margins and funds transfer prices

The goal of FTP is to lock the commercial margin rates $m_i^{(c)}(t,u)$ and $m_j^{(c)}(t,u)$ over the lifetime of the product contract

Commercial margin

Example

We consider the following interest earning and bearing items:

Assets	$N_i(t,u)$	$R_i(t,u)$	Liabilities	$N_j(t,u)$	$R_j(t,u)$
Loans	100	5%	Deposits	100	0.5%
Mortgages	100	4%	Debts	60	2.5%

The FTP for the loans and the mortgages is equal to 3%, while the FTP for deposits is equal to 1.5% and the FTP for debts is equal to 2.5%. We assume that the market rate is equal to 2.5%

Solution

We obtain the following results:

Assets	$m_i^{(c)}(t,u)$	$m_i^{(t)}(t,u)$	Liabilities	$m_j^{(c)}(t,u)$	$m_j^{(t)}(t,u)$
Loans	2%	0.5%	Deposits	1.0%	1.0%
Mortgages	1%	0.5%	Debts	0.0%	0.0%

Commercial margin

The commercial margin of the bank is equal to:

$$M^{(c)} = 100 \times 2\% + 100 \times 1\% + 100 \times 1\% + 60 \times 0\% = 4$$

For the transformation margin, we have:

$$M^{(t)} = 100 \times 0.5\% + 100 \times 0.5\% + 100 \times 1.0\% + 60 \times 0\% = 2$$

We don't have $M^{(c)} + M^{(t)} = \mathrm{NII}$ because assets and liabilities are not compensated:

$$\mathrm{NII} - \left(M^{(c)} + M^{(t)}\right) = \left(\mathrm{RA}\left(t,u\right) - \mathrm{RL}\left(t,u\right)\right) \cdot r\left(t,u\right) = 40 \times 2.5\% = 1$$

The commercial margin of each product is:

- $M^{(c)}_{Loans} = 2$
- $M^{(c)}_{Mortgages} = 1$
- $M^{(c)}_{Deposits} = 1$

Computing the internal transfer rate

The reference rate

- Since we have $m_i^{(t)}(t,u) = \mathrm{FTP}_i(t,u) - r(t,u)$, internal prices are fair if the corresponding mark-to-market is equal to zero on average
- For a contract with a bullet maturity, this implies that:

$$\mathrm{FTP}_i(t,u) = \mathbb{E}[r(t,u)]$$

- The transformation margin can then be interpreted as an interest rate swap receiving a fixed leg $\mathrm{FTP}_i(t,u)$ and paying a floating leg $r(t,u)$

Remark

It follows that the funds transfer price is equal to the market swap rate at the initial date t with the same maturity than the asset item i

Computing the internal transfer rate

FTPs and the new production

If we assume that the commercial margin rate of the business unit is constant:

$$R(u) - \mathrm{FTP}(t,u) = m$$

we can show that:

$$\mathrm{FTP}(t,u) = R(u) + \frac{\mathbb{E}_t\left[\int_t^\infty B(t,u)\,\mathbf{S}(t,u)\,(r(u) - R(u))\,\mathrm{d}u\right]}{\mathbb{E}_t\left[\int_t^\infty B(t,u)\,\mathbf{S}(t,u)\,\mathrm{d}u\right]}$$

We deduce that:

- for a loan with a fixed rate, the funds transfer price is exactly the swap rate with the same maturity than the loan and the same amortization scheme than the new production
- if the client rate $R(u)$ is equal to the short-term market rate $r(u)$, the funds transfer price $\mathrm{FTP}(t,u)$ is also equal to $r(u)$

Non-maturity deposits
What is the maturity of NMDs?

- The deposit balance of the client A is equal to \$500 $\Rightarrow$ the duration of this deposit is equal to zero day
- We consider 1 000 clients $\Rightarrow$ the total amount that may be withdrawn today is then between \$0 and \$500 000
- We assume that the probability to withdraw \$500 at once is equal to 50% $\Rightarrow$ the probability that \$500 000 are withdrawn is less than $10^{-300}\%$!
- Since we have $\Pr\{S > 275000\} < 0.1\%$, we can decide that 55% of the deposit balance has a duration of zero day, 24.75% has a duration of one day, 11.14% has a duration of two days, etc.

The statistical duration of NMDs is long (and not short)

Non-maturity deposits

Dynamic modeling

In the case of non-maturity deposits, the hazard function rate $\lambda(t, u)$ of the amortization function $\mathbf{S}(t, u)$ does not depend on the entry date t:

$$\lambda(t, u) = \lambda(u)$$

We can show that (HFRM, pages 428-428):

$$\mathrm{d}N(t) = (\mathrm{NP}(t) - \lambda(t) N(t))\, \mathrm{d}t$$

If we assume that the new production and the hazard rate are constant – $\mathrm{NP}(t) = \mathrm{NP}$ and $\lambda(t) = \lambda$, we obtain:

$$\mathrm{d}N(t) = \lambda(N_\infty - N(t))\, \mathrm{d}t$$

Non-maturity deposits
Dynamic modeling

Two extensions:

- The Ornstein-Uhlenbeck model:

$$\mathrm{d}N(t) = \lambda \left(N_\infty - N(t)\right) \mathrm{d}t + \sigma \,\mathrm{d}W(t)$$

- The aggregate model:

$$D(t) = \underbrace{D_\infty e^{g(t-s)}}_{D_{\mathrm{long}}(s,t)} + \underbrace{\left(D_s - D_\infty\right) e^{(g-\lambda)(t-s)} + \varepsilon(t)}_{D_{\mathrm{short}}(s,t)}$$

where g is the growth rate of deposits

Non-maturity deposits

Stable vs non-stable deposits

Remark

We have:

$$D(t) = \underbrace{\varphi D_\infty e^{g(t-s)}}_{D_{\text{stable}}(s,t)} + \underbrace{(D_s - D_\infty) e^{(g-\lambda)(t-s)} + \varepsilon(t) + (1-\varphi) D_\infty e^{g(t-s)}}_{D_{\text{non-stable}}(s,t)}$$

Calibration of φ

We have:

$$\Pr\{D(t) \leq \varphi D_\infty\} = 1 - \alpha$$

IIf we consider the Ornstein-Uhlenbeck dynamics, we obtain:

$$\varphi = 1 - \frac{\sigma \Phi^{-1}(1-\alpha)}{D_\infty \sqrt{2\lambda}}$$

Non-maturity deposits
Stable vs non-stable deposits

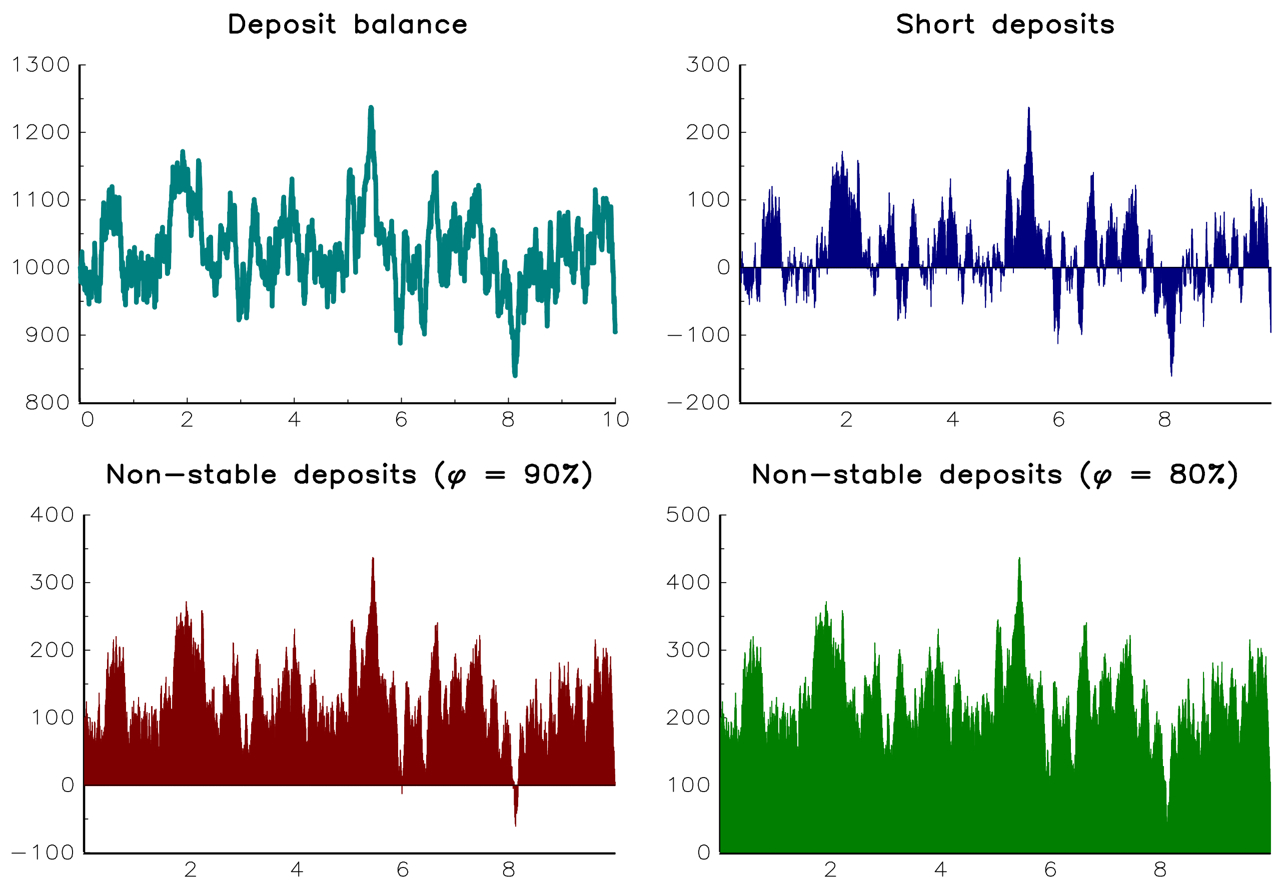


Figure: Stable and non-stable deposits

Non-maturity deposits

Behavioral modeling

The Hutchison-Pennacchi-Selvaggio framework

- The deposit rate $i(t)$ is exogenous and the bank account holder modifies his current deposit balance $D(t)$ to target a level $D^{\star}(t)$:

$$\ln D^{\star}(t) = \beta_0 + \beta_1 \ln i(t) + \beta_2 \ln Y(t)$$

  where $Y(t)$ is the income of the account holder

- The behavior of the bank account holder can be represented by a mean-reverting AR(1) process:

$$\ln D(t) - \ln D(t-1) = (1-\phi)\left(\ln D^{\star}(t) - \ln D(t-1)\right) + \varepsilon(t)$$

- The bank maximizes its profit $i^{\star}(t) = \arg\max \Pi(t)$ where the profit $\Pi(t)$ is equal to the revenue minus the cost:

$$\Pi(t) = r(t) \cdot D(t) - (i(t) + c(t)) \cdot D(t)$$

  $r(t)$ is the market interest rate and $c(t)$ is the cost of issuing deposits

- We can show that:

$$i^{\star}(t) = r(t) - s(t)$$

Non-maturity deposits
Behavioral modeling

The IRS framework (Jarrow and van Deventer, 1998)

- The current market value of deposits is the net present value of the cash flow stream $D(t)$:

$$V(0) = \mathbb{E}\left[\sum_{t=0}^{\infty} B(0, t+1)(r(t) - i(t)) D(t)\right]$$

$V(0)$ as an exotic interest rate swap, where the bank receives the market rate and pays the deposit rate.

- We have:

$$\ln D(t) = \ln D(t-1) + \beta_0 + \beta_1 r(t) + \beta_2 (r(t) - r(t-1)) + \beta_3 t$$

and:

$$i(t) = i(t) + \beta_0' + \beta_1' r(t) + \beta_2' (r(t) - r(t-1))$$

Non-maturity deposits

Behavioral modeling

Asymmetric adjustment models

- O'Brien model:

$$\Delta i(t) = \alpha(t) \cdot (\hat{\imath}(t) - i(t-1)) + \eta(t)$$

where $\hat{\imath}(t)$ is the conditional equilibrium deposit rate and:

$$\alpha(t) = \alpha^{+} \cdot \mathbb{1}\{\hat{\imath}(t) > i(t-1)\} + \alpha^{-} \cdot \mathbb{1}\{\hat{\imath}(t) < i(t-1)\}$$

- Frachot model:

$$\ln D(t) - \ln D(t-1) = (1-\phi)(\ln D^{\star}(t) - \ln D_{t-1}) + \delta_c(r(t), r^{\star})$$

where $\delta_c(r(t), r^{\star}) = \delta \cdot \mathbb{1}\{r(t) \leq r^{\star}\}$ and $r^{\star}$ is the interest rate floor

- OTS model:

$$d(t) = d(t-1) + \Delta \ln\left(\beta_0 + \beta_1 \arctan\left(\beta_2 + \beta_3 \frac{i(t)}{r(t)}\right) + \beta_4 i(t)\right) + \varepsilon(t)$$

Non-maturity deposits

Behavioral modeling

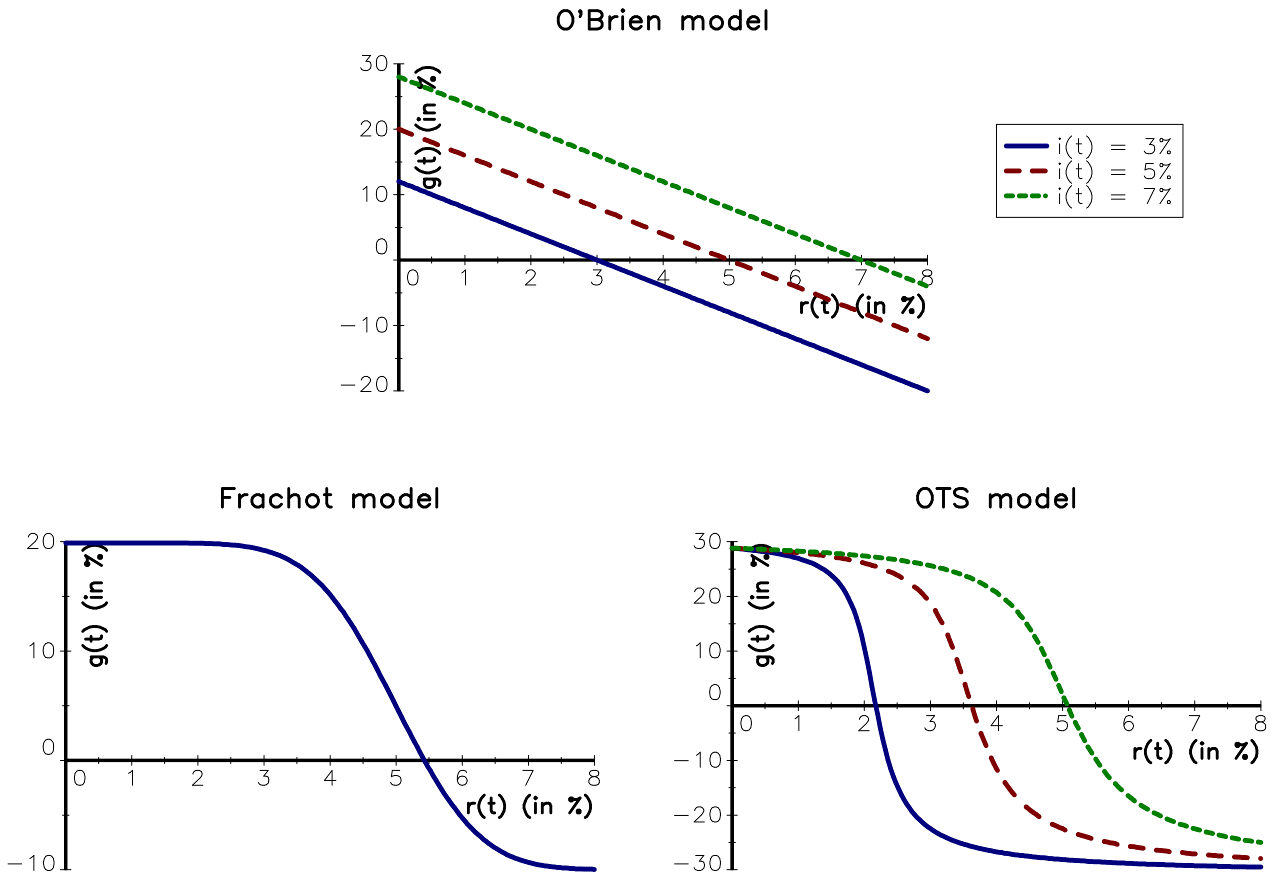


Figure: Impact of the market rate on the growth rate of deposits

Prepayment risk

Definition

A prepayment is the settlement of a debt or the partial repayment of its outstanding amount before its maturity date

Prepayment risk

Factors of prepayment

1. Refinancing:

$$\mathbb{P}(t) = \Pr\{\boldsymbol{\tau} \leq t\} = \vartheta(i_0 - i(t))$$

> *"A household with a 30-year fixed-rate mortgage of* \$200 000 *at an interest rate of* 6.0% *that refinances when rates fall to* 4.5% *(approximately the average rate decrease between 2008 and 2010 in the US) saves more than* \$60 000 *in interest payments over the life of the loan, even after accounting for refinance transaction costs. Further, when mortgage rates reached all-time lows in late 2012, with rates of roughly* 3.35% *prevailing for three straight months, this household with a contract rate of* 6.5% *would save roughly* \$130 000 *over the life of the loan by refinancing" (Keys, et al., 2016, pages 482-483).*

2. Housing turnover (marriage, divorce, death, children leaving home or changing jobs)

Prepayment risk
Structural models

The prepayment value is the premium of an American call option, meaning that we can derive the optimal option exercise. In this case, the prepayment strategy can be viewed as an arbitrage strategy between the market interest rate and the cost of refinancing

Prepayment risk

Reduced-form models

Rate, coupon or maturity incentive?

We assume that the mortgage rate drops from i_0 to $i(t)$

- The absolute difference of the annuity is equal to:

$$\mathfrak{D}_A(i_0, i(t)) = A(i_0, n) - A(i(t), n)$$

- The relative cumulative difference $\mathfrak{C}(i_0, i(t))$ is equal to:

$$\mathfrak{C}(i_0, i(t)) = \frac{\sum_{t=1}^{n} \mathfrak{D}_A(i_0, i(t))}{N_0}$$

- By assuming that the borrower continues to pay the same annuity, the maturity reduction is given by:

$$\mathfrak{N}(i_0, i(t)) = \{x \in \mathbb{N} : A(i(t), x) \geq A(i(t), n), A(i(t), x+1) < A(i(t), n)\}$$

Prepayment risk

Reduced-form models

Table: Impact of a new mortgage rate (100 KUSD, 5%, 10-year)

i (in %)	A (in \$)	$\mathfrak{D}_A$ (in \$) Monthly	Annually	$\mathfrak{D}_R$ (in %)	$\mathfrak{C}$ (in %)	$\mathfrak{M}$ (in years)
5.0	1 061					
4.5	1 036	24	291	2.3	2.9	9.67
4.0	1 012	48	578	4.5	5.8	9.42
3.5	989	72	862	6.8	8.6	9.17
3.0	966	95	1 141	9.0	11.4	8.92
2.5	943	118	1 415	11.1	14.2	8.75
2.0	920	141	1 686	13.2	16.9	8.50
1.5	898	163	1 953	15.3	19.5	8.33
1.0	876	185	2 215	17.4	22.2	8.17
0.5	855	206	2 474	19.4	24.7	8.00

Prepayment risk

Reduced-form models

Table: Impact of a new mortgage rate (100 KUSD, 5%, 20-year)

i (in %)	A (in \$)	$\mathfrak{D}_A$ (in \$) Monthly	$\mathfrak{D}_A$ (in \$) Annually	$\mathfrak{D}_R$ (in %)	$\mathfrak{C}$ (in %)	$\mathfrak{M}$ (in years)
5.0	660					
4.5	633	27	328	4.1	6.6	18.67
4.0	606	54	648	8.2	13.0	17.58
3.5	580	80	960	12.1	19.2	16.67
3.0	555	105	1 264	16.0	25.3	15.83
2.5	530	130	1 561	19.7	31.2	15.17
2.0	506	154	1 849	23.3	37.0	14.50
1.5	483	177	2 129	26.9	42.6	14.00
1.0	460	200	2 401	30.3	48.0	13.50
0.5	438	222	2 664	33.6	53.3	13.00

Prepayment risk

Reduced-form models

Table: Impact of a new mortgage rate (100 KUSD, 10%, 10-year)

i (in %)	A (in \$)	$\mathfrak{D}_A$ (in \$) Monthly	Annually	$\mathfrak{D}_R$ (in %)	$\mathfrak{C}$ (in %)	$\mathfrak{M}$ (in years)
10.0	1 322					
9.0	1 267	55	657	4.1	6.6	9.33
8.0	1 213	108	1 299	8.2	13.0	8.75
7.0	1 161	160	1 925	12.1	19.3	8.33
6.0	1 110	211	2 536	16.0	25.4	7.92
5.0	1 061	261	3 130	19.7	31.3	7.58
4.0	1 012	309	3 709	23.3	37.1	7.25
3.0	966	356	4 271	26.9	42.7	6.92
2.0	920	401	4 816	30.4	48.2	6.67
1.0	876	445	5 346	33.7	53.5	6.50

Prepayment risk

Reduced-form models

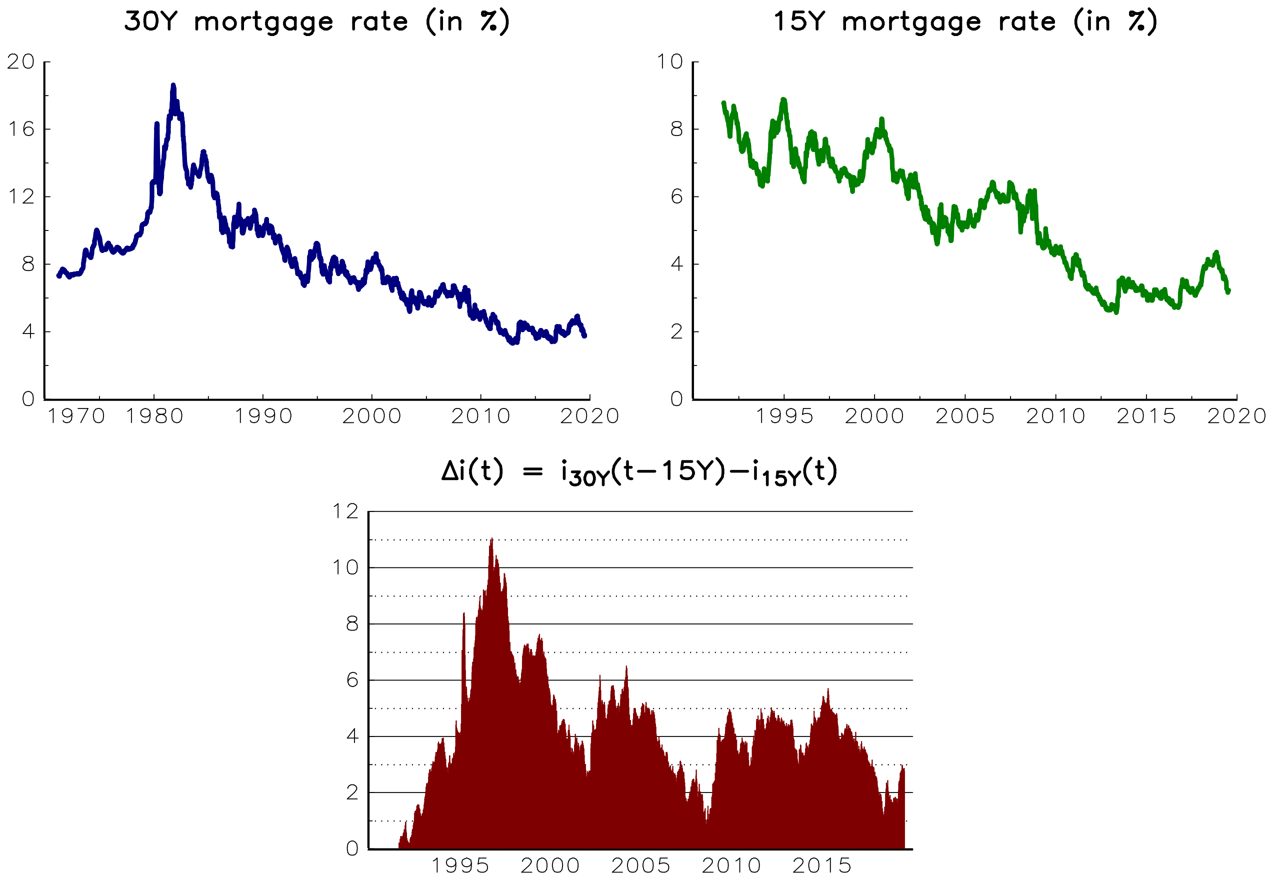


Figure: Evolution of 30-year and 15-year mortgage rates in the US

Prepayment risk

Survival function with prepayment risk

We have:

$$\mathbf{S}(t,u) = \mathbf{S}_c(t,u) \cdot \mathbf{S}_p(t,u)$$

where $\mathbf{S}_c(t,u)$ is the traditional amortization function (or the contract-based survival function) and $\mathbf{S}_p(t,u)$ is the prepayment-based survival function

Prepayment risk

Survival function with prepayment risk

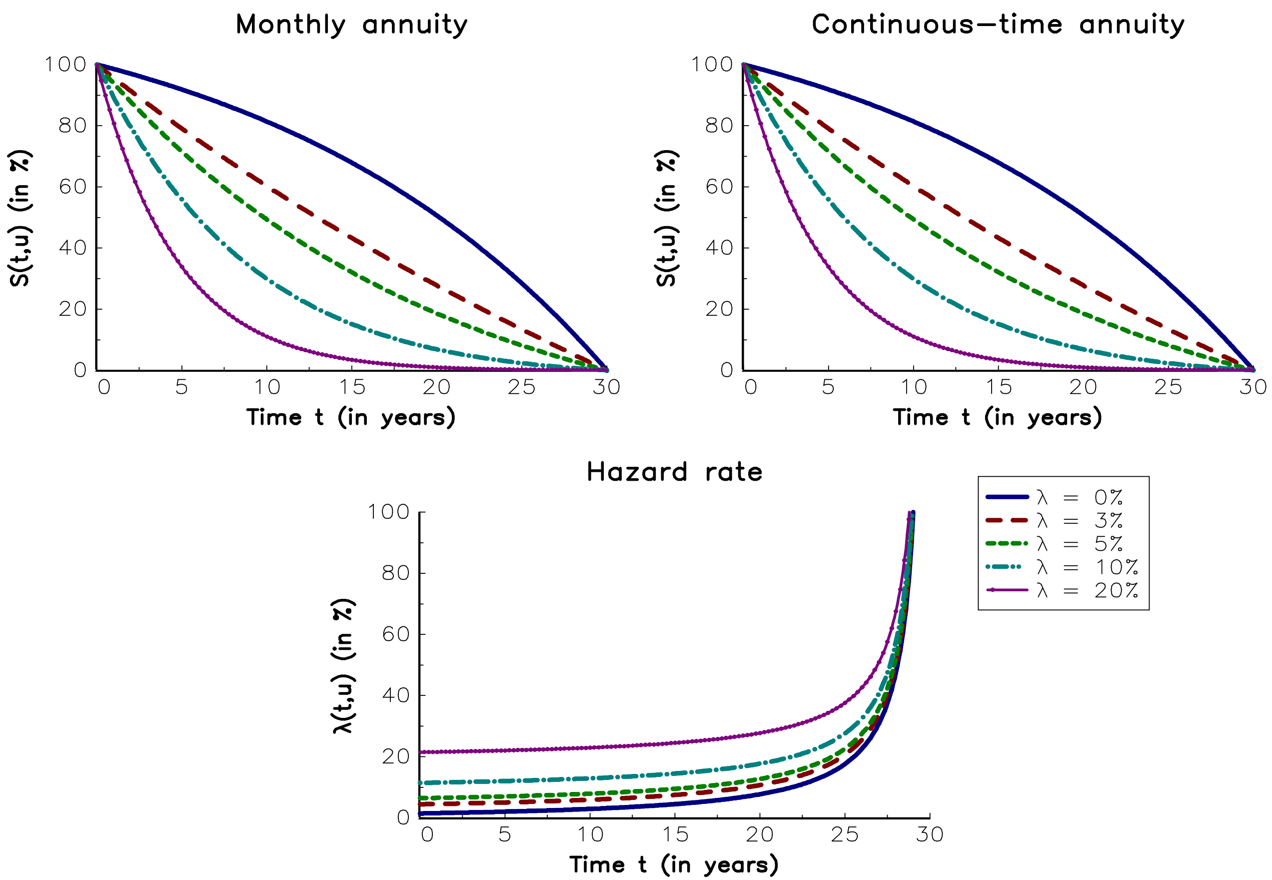


Figure: Survival function in the case of prepayment

Prepayment risk

Specification of the hazard function

- $\mathbf{S}_p(t,u)$ can be decomposed into the product of two survival functions:

$$\mathbf{S}_p(t,u) = \mathbf{S}_{\mathrm{refinancing}}(t,u) \cdot \mathbf{S}_{\mathrm{turnover}}(t,u)$$

- OTC model:

$$\lambda_p(t,u) = \lambda_{\mathrm{age}}(u-t) \cdot \lambda_{\mathrm{seasonality}}(u) \cdot \lambda_{\mathrm{rate}}(u)$$

where λ_{age} measures the impact of the loan age, $\lambda_{\mathrm{seasonality}}$ corresponds to the seasonality factor and λ_{rate} represents the influence of market rates

Prepayment risk

Specification of the hazard function

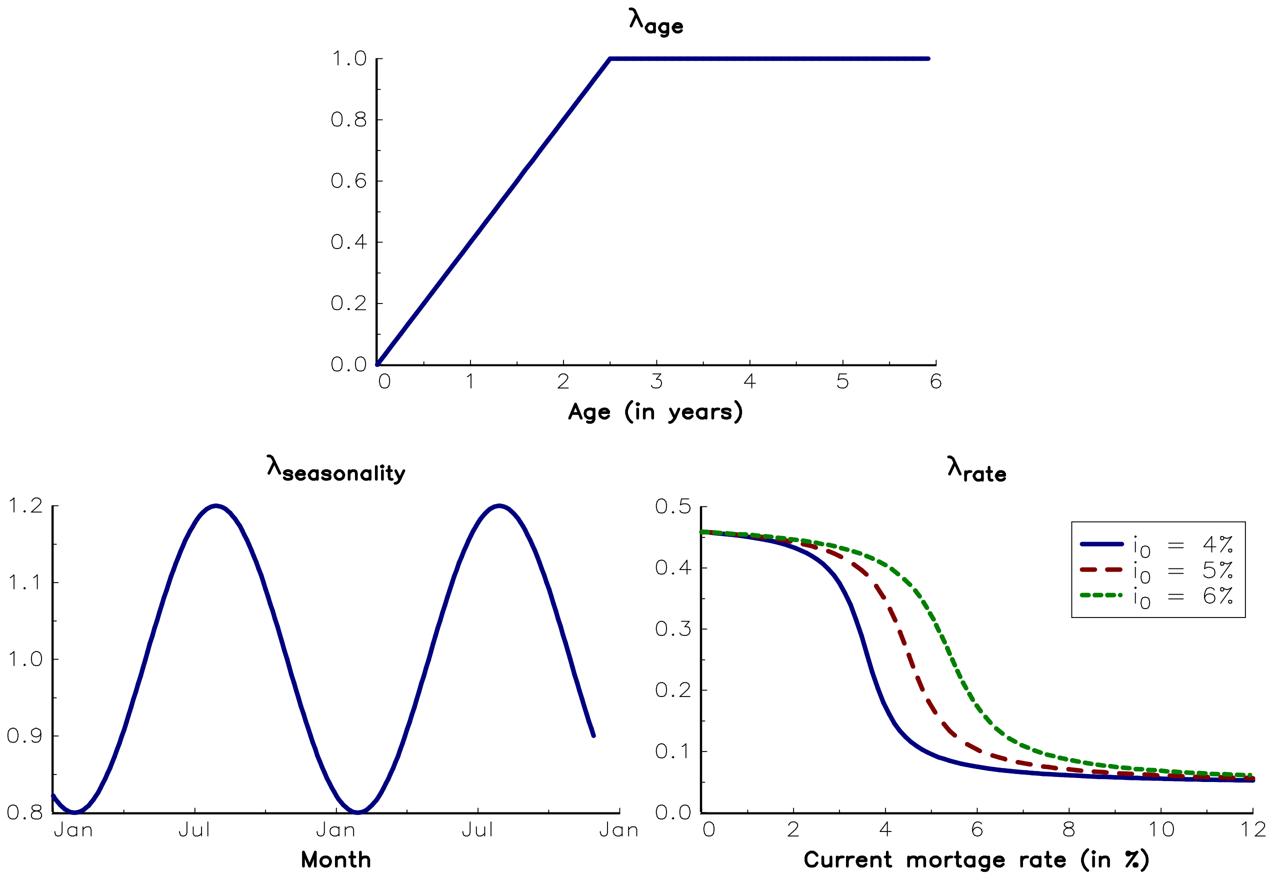


Figure: Components of the OTC model

Prepayment risk

Statistical measure of prepayment

- Single monthly mortality:

$$\mathrm{SMM} = \frac{\text{prepayments during the month}}{\text{outstanding amount at the beginning of the month}}$$

- The constant prepayment rate (CPR) and the SMM are related by the following equation:

$$\mathrm{CPR} = (1 - (1 - \mathrm{SMM}))^{12}$$

- In IRRBB, the CPR is also known as the conditional prepayment rate:

$$\begin{aligned} \mathrm{CPR}(u, t) &= \Pr\{u < \boldsymbol{\tau} \leq u + 1 \mid \boldsymbol{\tau} \geq u\} \\ &= \frac{\mathbf{S}_p(t, u) - \mathbf{S}_p(t, u + 1)}{\mathbf{S}_p(t, u)} \\ &= 1 - \exp\left(-\int_u^{u+1} \lambda_p(t, s)\, \mathrm{d}s\right) \end{aligned}$$

Prepayment risk

Statistical measure of prepayment

Table: Conditional prepayment rates in June 2018 by coupon rate and issuance date

Year	2012	2013	2014	2015	2016	2017	2018
Coupon = 3%	9.6%	10.2%	10.9%	10.0%	8.7%	5.3%	3.1%
Coupon = 4.5%	16.1%	15.8%	16.6%	17.9%	17.4%	12.8%	5.3%
Difference	6.5%	5.6%	5.7%	8.0%	8.7%	7.6%	2.2%

Redemption risk

The funding risk of term deposits

- A term deposit, also known as time deposit or certificate of deposit (CD), is a fixed-term cash investment. The client deposits a minimum sum of money into a banking account in exchange for a fixed rate over a specified period
- When buying a term deposit, the investor can withdraw their funds only after the term ends
- Under some conditions, the investor may withdraw his term deposit before the maturity date if he pays early redemption costs and fees

Redemption risk

Early time deposit withdrawals may be motivated by two reasons:

1. Economic motivation: $i(t) \gg i_0$
2. Negative liquidity shocks of depositors

The redemption-based survival function of time deposits can be decomposed as:

$$\mathbf{S}_r(t,u) = \mathbf{S}_{\text{economic}}(t,u) \cdot \mathbf{S}_{\text{liquidity}}(t,u)$$

Redemption risk

Modeling the early withdrawal risk

Early withdrawals due to economic reasons

- We note t the current date, m the maturity of the time deposit and N_0 the initial investment at time 0
- The value of the time deposit at the maturity is equal to $V_0 = N_0 (1 + i_0)^m$
- The value of the investment for $\boldsymbol{\tau} = t$ becomes:

$$V_r(t) = N_0 \cdot (1 + (1 - \varphi(t))\, i_0)^t \cdot (1 + i(t))^{m-t} - C(t)$$

  where $\varphi(t)$ is the penalty parameter applied to interest paid and $C(t)$ is the break fee
- The rational investor redeems the term deposit if the refinancing incentive is positive:

$$\mathrm{RI}(t) = \frac{V_r(t) - V_0}{N_0} > 0$$

Redemption risk

Modeling the early withdrawal risk

Early withdrawals due to economic reasons

- We can assume that:

$$\lambda_{\mathrm{economic}}(t, u) = g(i(u) - i_0)$$

or:

$$\lambda_{\mathrm{economic}}(t, u) = g(r(u) - i_0)$$

Redemption risk

Modeling the early withdrawal risk

Early withdrawals due to negative liquidity shocks

We can decompose the hazard function into two effects:

$$\lambda_{\mathrm{liquidity}}\left(t,u\right)=\lambda_{\mathrm{structural}}+\lambda_{\mathrm{cyclical}}\left(u\right)$$

where $\lambda_{\mathrm{structural}}$ is the structural rate of redemption and $\lambda_{\mathrm{cyclical}}\left(u\right)$ is the liquidity component due to the economic cycle. A simple way to model $\lambda_{\mathrm{cyclical}}\left(u\right)$ is to consider a linear function of the GDP growth

Exercises

- Interest rate risk
  - Exercise 7.4.1 – Constant amortization of a loan
  - Exercise 7.4.2 – Computation of the amortization functions $\mathbf{S}(t,u)$ and $\mathbf{S}^{\star}(t,u)$
  - Exercise 7.4.3 – Continuous-time analysis of the constant amortization mortgage (CAM)
- Non-maturity deposits (NMD)
  - Exercise 7.4.4 – Valuation of non-maturity deposits
- Prepayment risk
  - Exercise 7.4.5 – Impact of prepayment on the amortization scheme of the CAM

References

Basel Committee on Banking Supervision (2016)
Interest Rate Risk in the Banking Book, April 2016.

Bessis, J. (2015)
Risk Management in Banking, Fourth edition, John Wiley & Sons.

Demey, P., Frachot, A., and Riboulet, G. (2003)
Introduction à la Gestion Actif-Passif Bancaire, Economica.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 7.

Roncalli, T. (2020)
Handbook of Financial Risk Management – Companion Book, Chapter 7.

Course 2023-2024 in Financial Risk Management
Lecture 8. Model Risk

Thierry Roncalli⋆

⋆Amundi Asset Management[17]

⋆University of Paris-Saclay

September 2023

[17]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- **Lecture 8: Model Risk**
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Main issue

Easy

Computing an option price using a stochastic model

Hard

Computing an option price that corresponds to the cost of the hedging strategy

$$\text{Pricing} = \text{Hedging}$$

The Black-Scholes model

Black and Scholes (1973) assumed that the dynamics of the asset price $S(t)$ is given by a GBM:

$$\begin{cases} \mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}W(t) \\ S(t_0) = S_0 \end{cases}$$

where:

- S_0 is the current price
- μ is the drift
- σ is the volatility of the diffusion
- $W(t)$ is a standard Brownian motion

The Black-Scholes model

Let $f(S(T))$ be the payoff of a contingent claim where T is the maturity of the derivative contract. The price V of the contingent claim is then equal to the cost of the hedging portfolio. We can show that:

$$\begin{cases} \frac{1}{2}\sigma^2 S^2 \partial_S^2 V(t,S) + (\mu - \lambda(t)\sigma) S \partial_S V(t,S) + \partial_t V(t,S) - r(t) V(t,S) = 0 \\ V(T, S(T)) = f(S(T)) \end{cases}$$

This equation is called the fundamental pricing equation

The Black-Scholes model

The function $\lambda(t)$ is interpreted as the risk price of the Wiener process $W(t)$:

$$\lambda(t) = \frac{\mu - b(t)}{\sigma}$$

where $b(t)$ is the cost-of-carry

We have:

$$\begin{cases} \frac{1}{2}\sigma^2 S^2 \partial_S^2 V(t,S) + b(t) S \partial_S V(t,S) + \partial_t V(t,S) - r(t) V(t,S) = 0 \\ V(T, S(T)) = f(S(T)) \end{cases}$$

The current price of the derivatives contract is equal to $V(t_0, S_0)$

The Black-Scholes model

- Girsanov theorem with $g(t) = -\lambda(t)$:

$$\begin{cases} \mathrm{d}S(t) = b(t)\, S(t)\, \mathrm{d}t + \sigma S(t)\, \mathrm{d}W^{\mathbb{Q}}(t) \\ S(t_0) = S_0 \end{cases}$$

- $W^{\mathbb{Q}}(t)$ is a Brownian motion under the probability $\mathbb{Q}$ defined by:

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \exp\left(-\int_0^t \lambda(s)\, \mathrm{d}W(s) - \frac{1}{2}\int_0^t \lambda^2(s)\, \mathrm{d}s\right)$$

- Feynman-Kac formula with $h(t,x) = r(t)$ and $g(t,x) = 0$:

$$V_0 = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r(t)\, \mathrm{d}t} f(S(T)) \middle| \mathcal{F}_0 \right]$$

- V_0 is called the martingale solution and $\mathbb{Q}$ is called the risk-neutral probability measure

Application to European options

We consider an European call option whose payoff at maturity is equal to:

$$\mathcal{C}(T) = (S(T) - K)^{+}$$

We assume that the interest rate $r(t)$ and the cost-of-carry parameter $b(t)$ are constant:

$$\begin{aligned}
\mathcal{C}_0 &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r\,\mathrm{d}t} (S(T) - K)^{+} \middle| \mathcal{F}_0 \right] \\
&= e^{-rT}\mathbb{E}\left[\left(S_0 e^{\left(b - \frac{1}{2}\sigma^2\right)T + \sigma W^{Q}(T)} - K \right)^{+} \right] \\
&= e^{-rT} \int_{-d_2}^{\infty} \left(S_0 e^{\left(b - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}x} - K \right) \phi(x)\,\mathrm{d}x \\
&= S_0 e^{(b-r)T}\Phi(d_1) - K e^{-rT}\Phi(d_2)
\end{aligned}$$

where:

$$\begin{aligned}
d_1 &= \frac{1}{\sigma\sqrt{T}}\left(\ln\frac{S_0}{K} + bT \right) + \frac{1}{2}\sigma\sqrt{T} \\
d_2 &= d_1 - \sigma\sqrt{T}
\end{aligned}$$

Application to European options

Let us now consider an European put option with the following payoff:

$$\mathcal{P}(T) = (K - S(T))^{+}$$

We have:

$$\mathcal{C}(T) - \mathcal{P}(T) = (S(T) - K)^{+} - (K - S(T))^{+} = S(T) - K$$

We deduce that:

$$\begin{aligned} \mathcal{C}_0 - \mathcal{P}_0 &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r \,\mathrm{d}t} (S(T) - K) \middle| \mathcal{F}_0 \right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[e^{-rT} S(T) \middle| \mathcal{F}_0 \right] - K e^{-rT} \\ &= S_0 e^{(b-r)T} - K e^{-rT} \end{aligned}$$

This equation is known as the put-call parity and we have:

$$\begin{aligned} \mathcal{P}_0 &= \mathcal{C}_0 - S_0 e^{(b-r)T} + K e^{-rT} \\ &= -S_0 e^{(b-r)T} \Phi(-d_1) + K e^{-rT} \Phi(-d_2) \end{aligned}$$

Application to European options

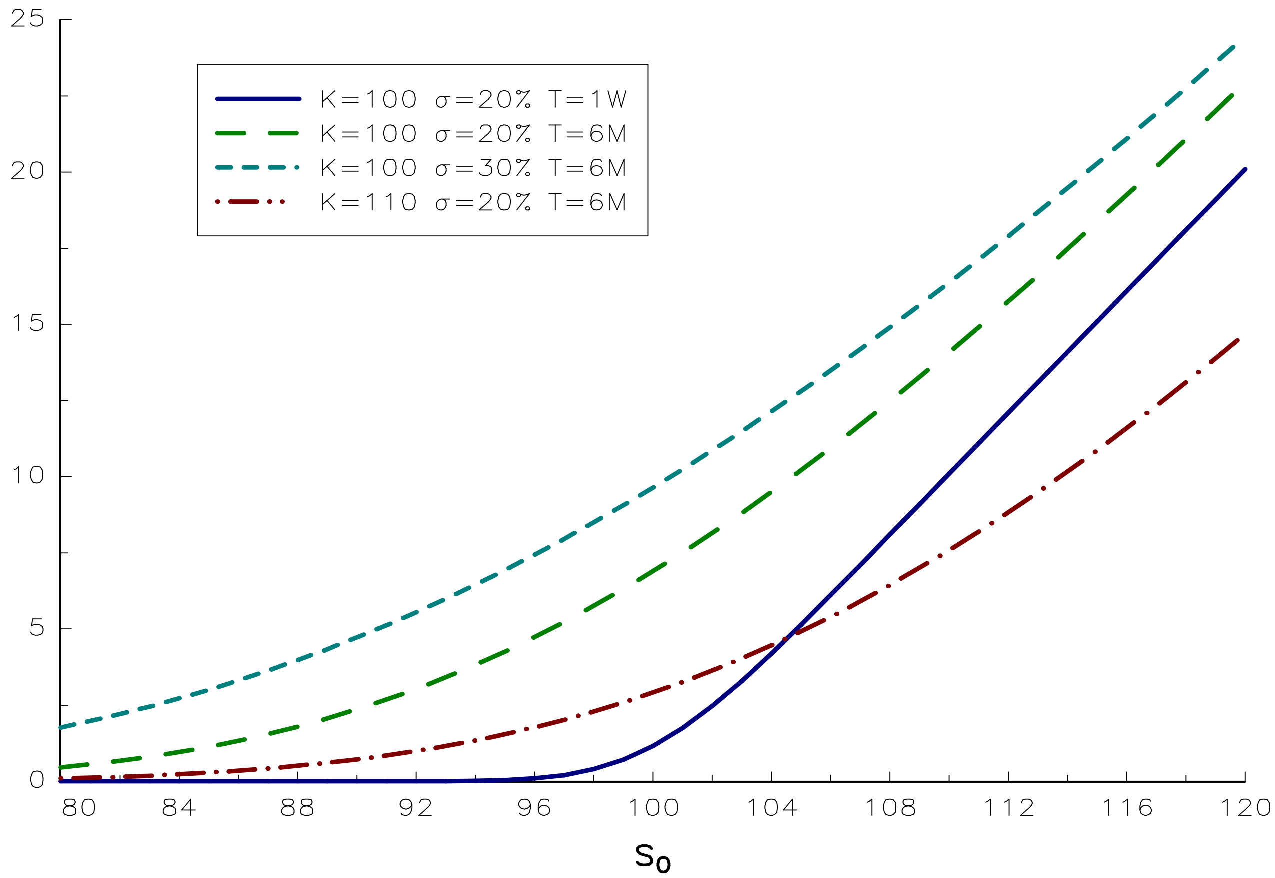


Figure: Price of the call option

Application to European options

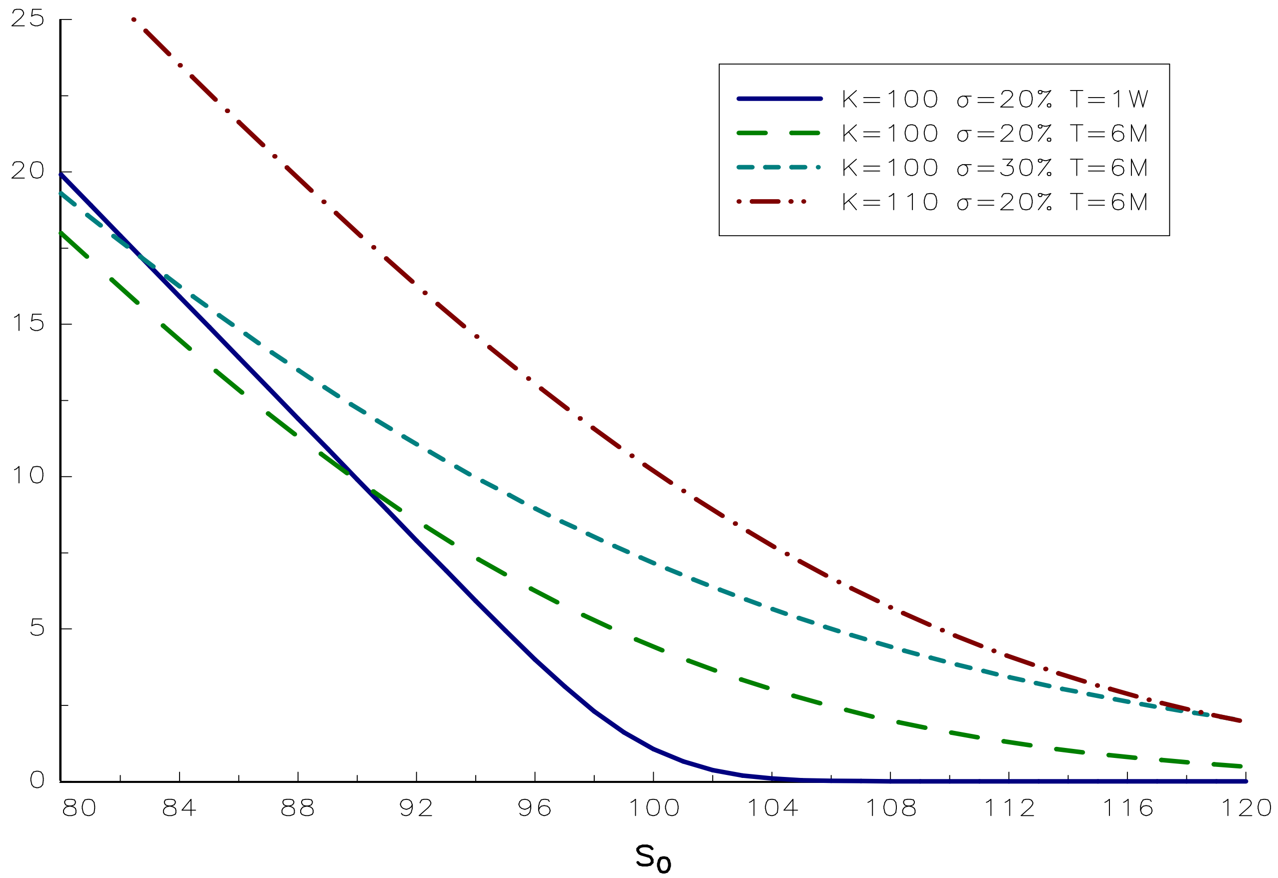


Figure: Price of the put option

Principle of dynamic hedging

- n assets that do not pay dividends or coupons during the period $[0, T]$
- For asset i, we have:

$$S_i(t) = S_i(0) + \int_0^t \mu_i(u)\,\mathrm{d}u + \int_0^t \sigma_i(u)\,\mathrm{d}W_i(u)$$

- We set up a trading portfolio $(\phi_1(t), \ldots, \phi_n(t))$ invested in the assets $(S_1(t), \ldots, S_n(t))$
- The value of this trading portfolio is:

$$X(t) = \sum_{i=1}^{n} \phi_i(t) S_i(t)$$

Self-financing strategy

- The portfolio is self-financing if the following conditions hold:

$$\left\{ \begin{array}{l} \mathrm{d}X(t) - \sum_{i=1}^{n} \phi_i(t)\,\mathrm{d}S_i(t) = 0 \\ X(0) = 0 \end{array} \right.$$

  1. The first condition means that all trades are financed by selling or buying assets in the portfolio
  2. The second condition implies that we don't need money to set up the initial portfolio

- This implies that:

$$\begin{array}{rcl} X(t) & = & X_0 + \sum_{i=1}^{n} \int_0^t \phi_i(u)\,\mathrm{d}S_i(u) \\ & = & \sum_{i=1}^{n} \phi_i(0)\,S_i(0) + \sum_{i=1}^{n} \int_0^t \phi_i(u)\,\mathrm{d}S_i(u) \end{array}$$

Principle of dynamic hedging

- We have:

$$\mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}W(t)$$

- The risk-free asset $B(t)$ satisfies:

$$\mathrm{d}B(t) = rB(t)\,\mathrm{d}t$$

- We set up a trading portfolio $(\phi(t), \psi(t))$ invested in the stock $S(t)$ and the risk-free asset $B(t)$
- The value of this portfolio is:

$$V(t) = \phi(t) S(t) + \psi(t) B(t)$$

Application to the Black-Scholes model

- We form a strategy $X(t)$ in which we are long the call option $\mathcal{C}(t, S(t))$ and short the trading portfolio $V(t)$:

$$\begin{aligned} X(t) &= \mathcal{C}(t, S(t)) - V(t) \\ &= \mathcal{C}(t, S(t)) - \phi(t) S(t) - \psi(t) B(t) \end{aligned}$$

- We have:

$$\begin{aligned} \mathrm{d}X(t) &= \partial_S \mathcal{C}(t, S(t)) \, \mathrm{d}S(t) + \\ &\quad \left(\partial_t \mathcal{C}(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) \, \partial_S^2 \mathcal{C}(t, S(t))\right) \mathrm{d}t - \\ &\quad \phi(t) \, \mathrm{d}S(t) - \psi(t) \, \mathrm{d}B(t) \end{aligned}$$

- By assuming that $\phi(t) = \partial_S \mathcal{C}(t, S(t))$, we obtain:

$$\mathrm{d}X(t) = \left(\partial_t \mathcal{C}(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t) \, \partial_S^2 \mathcal{C}(t, S(t)) - r\psi(t) B(t)\right) \mathrm{d}t$$

Application to the Black-Scholes model

- $X(t)$ is self-financing if $\mathrm{d}X(t) = 0$ or:

$$\psi(t) = \frac{\partial_t \mathcal{C}(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)\, \partial_S^2 \mathcal{C}(t, S(t))}{rB(t)}$$

- We deduce that:

$$\begin{aligned} \mathcal{C}(t, S(t)) &= \phi(t) S(t) + \psi(t) B(t) \\ &= \partial_S \mathcal{C}(t, S(t)) S(t) + \\ &\quad \frac{\partial_t \mathcal{C}(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)\, \partial_S^2 \mathcal{C}(t, S(t))}{rB(t)} B(t) \end{aligned}$$

Application to the Black-Scholes model

- This implies that $\mathcal{C}(t, S(t))$ satisfies the following PDE:

$$\frac{1}{2}\sigma^2 S^2 \partial_S^2 \mathcal{C}(t, S) + rS\partial_S \mathcal{C}(t, S) + \partial_t \mathcal{C}(t, S) - r\mathcal{C}(t, S) = 0$$

- Since $X(t)$ is self-financing ($X(t) = 0$), we also deduce that the trading portfolio $V(t)$ is the replicating portfolio of the call option:

$$\begin{aligned} V(t) &= \phi(t) S(t) + \psi(t) B(t) \\ &= \mathcal{C}(t, S(t)) - X(t) \\ &= \mathcal{C}(t, S(t)) \end{aligned}$$

Application to the Black-Scholes model

- If we define the replicating cost as follows:

$$\begin{aligned} C(t) &= \int_0^t \phi(u)\,\mathrm{d}S(u) + \int_0^t \psi(u)\,\mathrm{d}B(u) \\ &= \int_0^t \left(\mu S(u)\,\phi(u) + rB(u)\,\psi(u)\right)\mathrm{d}u + \int_0^T \sigma S(u)\,\phi(u)\,\mathrm{d}W(u) \end{aligned}$$

we have:

$$\begin{aligned} C(t) &= \int_0^t \mu S(u)\,\partial_S \mathcal{C}(u, S(u))\,\mathrm{d}u + \int_0^T \sigma S(u)\,\partial_S \mathcal{C}(u, S(u))\,\mathrm{d}W(u) \\ & \int_0^t \left(\partial_t \mathcal{C}(u, S(u)) + \frac{1}{2}\sigma^2 S^2(u)\,\partial_S^2 \mathcal{C}(u, S(u))\right)\mathrm{d}u \\ &= \int_0^t \mathrm{d}\mathcal{C}(u, S(u)) = \mathcal{C}(t, S(t)) - \mathcal{C}(0, S_0) \end{aligned}$$

- We verify that the replicating cost is exactly equal to the P&L of the long exposure on the call option

Cost-of-carry

- Let us consider a stock that pays a continuous dividend δ, the self-financing portfolio is:

$$X(t) = \mathcal{C}(t, S(t)) - \phi(t) S(t) - \psi(t) B(t)$$

We deduce that the change in the value of this portfolio is:

$$\mathrm{d}X(t) = \mathrm{d}\mathcal{C}(t, S(t)) - \phi(t)\,\mathrm{d}S(t) - \psi(t)\,\mathrm{d}B(t) - \phi(t) \cdot \underbrace{\delta \cdot S(t)\,\mathrm{d}t}_{\text{dividend}}$$

- Using the same rationale than previously, we obtain $\phi(t) = \partial_S \mathcal{C}(t, S(t))$ and:

$$\psi(t) = \frac{\partial_t \mathcal{C}(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)\,\partial_S^2 \mathcal{C}(t, S(t)) - \delta S(t)\,\partial_S \mathcal{C}(t, S(t))}{rB(t)}$$

Finally, we obtain the following PDE:

$$\frac{1}{2}\sigma^2 S^2 \partial_S^2 \mathcal{C}(t, S) + (r - \delta) S \partial_S \mathcal{C}(t, S) + \partial_t \mathcal{C}(t, S) - r\mathcal{C}(t, S) = 0$$

Cost-of-carry

- When the stock does not pay a dividend, the cost-of-carry parameter b is equal to the interest rate r
- When the stock pays a continuous dividend, the cost-of-carry parameter b is equal to $r - \delta$
- In the case of futures or forward contracts, the cost-of-carry is equal to zero
- For currency options, the cost-of-carry is the difference between the domestic interest rate r and the foreign interest rate $r^{\star}$

Table: Impact of the dividend on the option premium

	Put option				Call option			
S_0 / δ	0.00	0.02	0.05	0.07	0.00	0.02	0.05	0.07
90	1.28	1.44	1.73	1.94	13.50	12.67	11.48	10.72
100	4.42	4.83	5.50	5.97	6.89	6.31	5.50	5.00
110	10.19	10.87	11.91	12.63	2.91	2.59	2.16	1.90

Delta hedging

- The Black-Scholes model assumes that the replicating portfolio is rebalanced continuously
- In practice, it is rebalanced at some fixed dates t_i:

$$0 = t_0 < t_1 < \cdots < t_n = T$$

- At the initial date, we have:

$$X(t_0) = \mathcal{C}(t_0, S(t_0)) - V(t_0) = 0$$

where:

$$V(t_0) = \phi(t_0) \cdot S(t_0) + \psi(t_0) \cdot B(t_0)$$

- Because we have $\phi(t_0) = \boldsymbol{\Delta}(t_0)$ and $X(t_0) = 0$, we deduce that:

$$\psi(t_0) = \mathcal{C}(t_0, S(t_0)) - \boldsymbol{\Delta}(t_0) S(t_0)$$

Delta hedging

- At time t_1, the value of the replicating portfolio is then equal to:

$$V(t_1) = \boldsymbol{\Delta}(t_0) S(t_1) + (\mathcal{C}(t_0, S(t_0)) - \boldsymbol{\Delta}(t_0) S(t_0)) \cdot (1 + r(t_0)(t_1 - t_0))$$

- It follows that:

$$X(t_1) = \mathcal{C}(t_1, S(t_1)) - V(t_1)$$

- We are not sure that $X(t_1) = 0$ because it is not possible to hedge the jump $S(t_1) - S(t_0)$. We rebalance the portfolio and we have:

$$V(t_1) = \phi(t_1) \cdot S(t_1) + \psi(t_1) \cdot B(t_1)$$

- We deduce that:

$$\phi(t_1) = \boldsymbol{\Delta}(t_1)$$

and:

$$\psi(t_1) = V(t_1) - \boldsymbol{\Delta}(t_1) S(t_1)$$

Delta hedging

- At time t_2, the value of the replicating portfolio is equal to:

$$V(t_2) = \boldsymbol{\Delta}(t_1) S(t_2) + (V(t_1) - \boldsymbol{\Delta}(t_1) S(t_1)) \cdot (1 + r(t_1)(t_2 - t_1))$$

- More generally, we have:

$$X(t_i) = \mathcal{C}(t_i, S(t_i)) - V(t_i)$$

and:

$$V(t_i) = \underbrace{\boldsymbol{\Delta}(t_{i-1}) S(t_i)}_{V_S(t_i)} + \underbrace{(V(t_{i-1}) - \boldsymbol{\Delta}(t_{i-1}) S(t_{i-1})) \cdot (1 + r(t_{i-1})(t_i - t_{i-1}))}_{V_B(t_i)}$$

where $V_S(t_i)$ is the component due to the delta exposure on the asset and $V_B(t_i)$ is the component due to the cash exposure on the risk-free bond

Delta hedging

- We notice that:

$$\begin{aligned} V_S(t_i) &= \boldsymbol{\Delta}(t_{i-1}) \cdot S(t_i) \\ &= \boldsymbol{\Delta}(t_{i-1}) \cdot S(t_{i-1}) \cdot (1 + R_S(t_{i-1}; t_i)) \end{aligned}$$

and:

$$\begin{aligned} V_B(t_i) &= (V(t_{i-1}) - \boldsymbol{\Delta}(t_{i-1}) \cdot S(t_{i-1})) \cdot (1 + r(t_{i-1}) \cdot (t_i - t_{i-1})) \\ &= (V(t_{i-1}) - \boldsymbol{\Delta}(t_{i-1}) \cdot S(t_{i-1})) \cdot (1 + R_B(t_{i-1}; t_i)) \end{aligned}$$

where $R_S(t_{i-1}; t_i)$ and $R_B(t_{i-1}; t_i)$ are the asset and bond returns between t_{i-1} and t_i

Delta hedging

- At the maturity, we obtain:

$$\begin{aligned} X(T) &= X(t_n) \\ &= (S(T) - K)^+ - V(t_n) \end{aligned}$$

- $\Pi(T) = -X(T)$ is the P&L of the delta hedging strategy. To measure its efficiency, we consider the ratio π defined as follows:

$$\pi = \frac{\Pi(T)}{\mathcal{C}(t_0, S(t_0))}$$

Delta hedging

Example #1

We consider the replication of 100 ATM call options. The current price of the asset is 100 and the maturity of the option is 20 weeks. We consider the following parameter: $b = r = 5\%$ and $\sigma = 20\%$. We rebalance the replicating portfolio every week.

Delta hedging

- $T = 20/52$
- $K = 100$
- $\mathcal{C}(t_0, S(t_0)) = \$5.90$
- The replicating portfolio is rebalanced at times t_i:

$$t_i = \frac{i}{52}$$

Delta hedging

Table: An example of delta hedging strategy (negative P&L)

i	t_i	$S(t_i)$	$\boldsymbol{\Delta}(t_{i-1})$	$V_S(t_i)$	$V_B(t_i)$	$V(t_i)$	$\mathcal{C}(t_i, S(t_i))$	$X(t_i)$	$\Pi(t_i)$
0	0.00	100.00	0.00	0.00	590.90	590.90	590.90	0.00	0.00
1	0.02	95.63	58.59	5603.15	−5273.36	329.79	350.22	20.43	−20.43
2	0.04	95.67	43.72	4182.80	−3854.96	327.84	336.15	8.31	−8.31
3	0.06	94.18	43.24	4072.36	−3812.62	259.75	260.57	0.82	−0.82
4	0.08	92.73	37.29	3457.72	−3255.16	202.55	196.22	−6.33	6.33
5	0.10	96.59	31.34	3027.23	−2706.31	320.93	326.47	5.54	−5.54
6	0.12	101.68	44.63	4537.99	−3993.73	544.26	582.71	38.45	−38.45
7	0.13	101.41	63.39	6428.19	−5906.72	521.47	545.64	24.17	−24.17
8	0.15	100.22	62.36	6249.97	−5808.29	441.68	453.62	11.94	−11.94
9	0.17	99.32	57.57	5718.25	−5333.51	384.74	382.58	−2.16	2.16
10	0.19	101.64	53.46	5433.52	−4929.49	504.03	495.99	−8.04	8.04
11	0.21	101.81	63.27	6441.30	−5932.22	509.08	483.87	−25.21	25.21
12	0.23	102.62	64.10	6578.19	−6022.97	555.22	513.53	−41.69	41.69
13	0.25	107.56	67.97	7311.26	−6426.42	884.84	876.68	−8.16	8.16
14	0.27	102.05	86.90	8867.94	−8470.05	397.89	424.07	26.18	−26.18
15	0.29	100.88	66.19	6677.01	−6362.67	314.34	321.76	7.41	−7.41
16	0.31	106.90	59.86	6399.37	−5730.15	669.21	756.02	86.80	−86.80
17	0.33	107.66	90.32	9723.75	−8994.54	729.22	806.47	77.25	−77.25
18	0.35	101.79	94.74	9643.97	−9480.00	163.96	276.24	112.27	−112.27
19	0.37	101.76	69.88	7111.04	−6955.85	155.19	228.08	72.89	−72.89
20	0.38	101.83	75.10	7647.28	−7494.04	153.24	183.00	29.76	**−29.76**

Delta hedging

Table: An example of delta hedging strategy (positive P&L)

i	t_i	$S(t_i)$	$\Delta(t_{i-1})$	$V_S(t_i)$	$V_B(t_i)$	$V(t_i)$	$\mathcal{C}(t_i, S(t_i))$	$X(t_i)$	$\Pi(t_i)$
0	0.00	100.00	0.00	0.00	590.90	590.90	590.90	0.00	0.00
1	0.02	98.50	58.59	5771.31	−5273.36	497.95	489.70	−8.25	8.25
2	0.04	97.00	53.45	5184.51	−4771.31	413.19	396.75	−16.44	16.44
3	0.06	95.47	47.89	4571.99	−4236.14	335.85	311.62	−24.24	24.24
4	0.08	98.17	41.87	4110.19	−3664.81	445.38	419.94	−25.44	25.44
5	0.10	100.48	51.10	5134.88	−4575.85	559.03	528.68	−30.35	30.35
6	0.12	102.92	59.19	6092.33	−5394.04	698.28	664.00	−34.29	34.29
7	0.13	105.50	67.69	7140.94	−6274.05	866.89	829.99	−36.90	36.90
8	0.15	101.81	76.13	7750.53	−7171.44	579.09	550.21	−28.88	28.88
9	0.17	100.65	63.86	6427.97	−5928.66	499.31	457.48	−41.83	41.83
10	0.19	98.86	59.15	5847.59	−5459.40	388.19	337.04	−51.15	51.15
11	0.21	99.26	50.91	5053.11	−4649.03	404.09	335.31	−68.78	68.78
12	0.23	101.78	52.25	5317.65	−4786.50	531.15	458.03	−73.12	73.12
13	0.25	99.28	64.14	6367.78	−6002.74	365.03	288.19	−76.84	76.84
14	0.27	99.19	51.19	5077.96	−4722.07	355.89	257.52	−98.36	98.36
15	0.29	95.53	49.97	4773.36	−4604.77	168.59	92.40	−76.18	76.18
16	0.31	98.02	26.47	2594.85	−2362.61	232.23	148.05	−84.19	84.19
17	0.33	97.03	39.61	3843.35	−3653.84	189.51	83.97	−105.54	105.54
18	0.35	96.64	29.34	2835.17	−2659.65	175.51	44.51	−131.01	131.01
19	0.37	95.01	21.11	2005.37	−1866.05	139.32	3.75	−135.56	135.56
20	0.38	93.67	3.62	338.73	−204.45	134.27	0.00	−134.27	**134.27**

Delta hedging

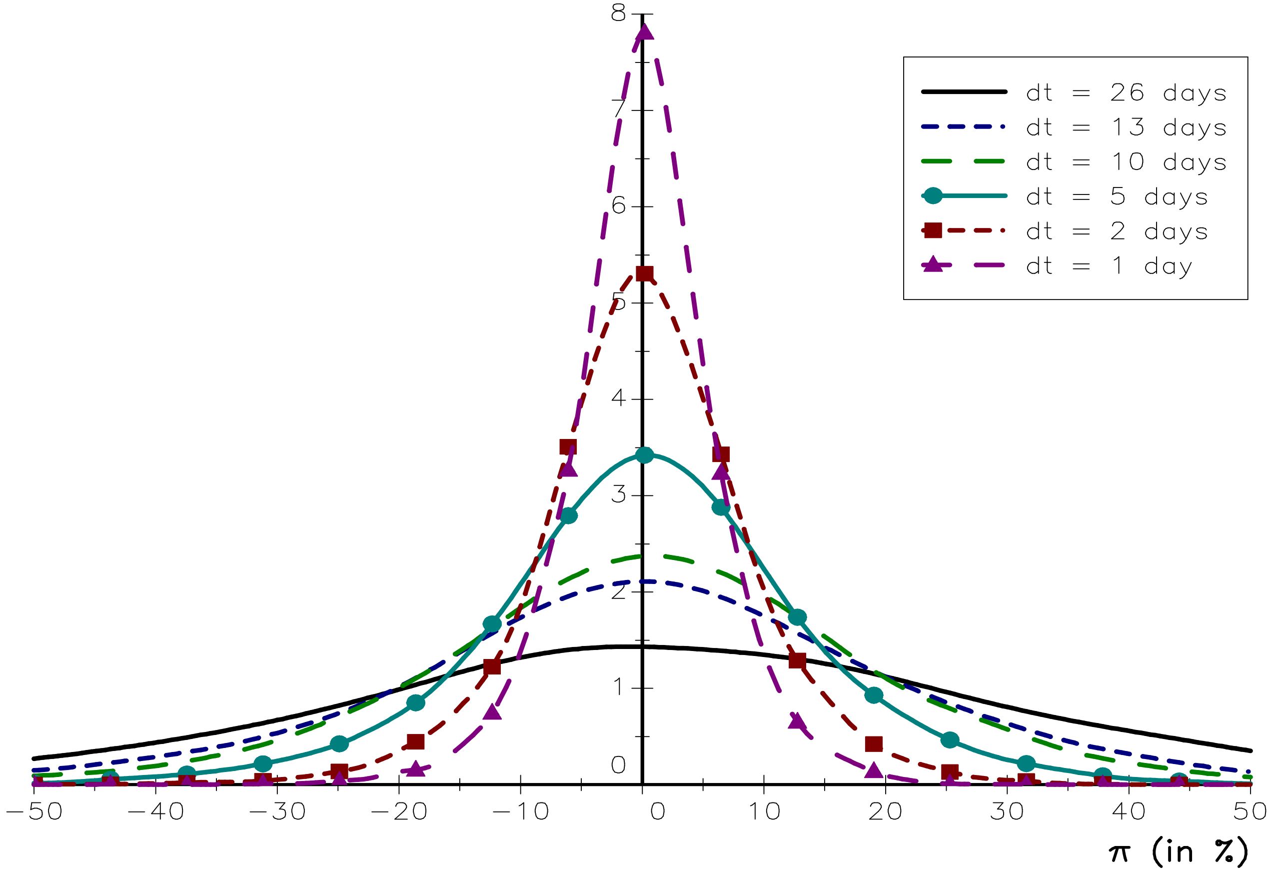


Figure: Probability density function of the hedging ratio π

Delta hedging

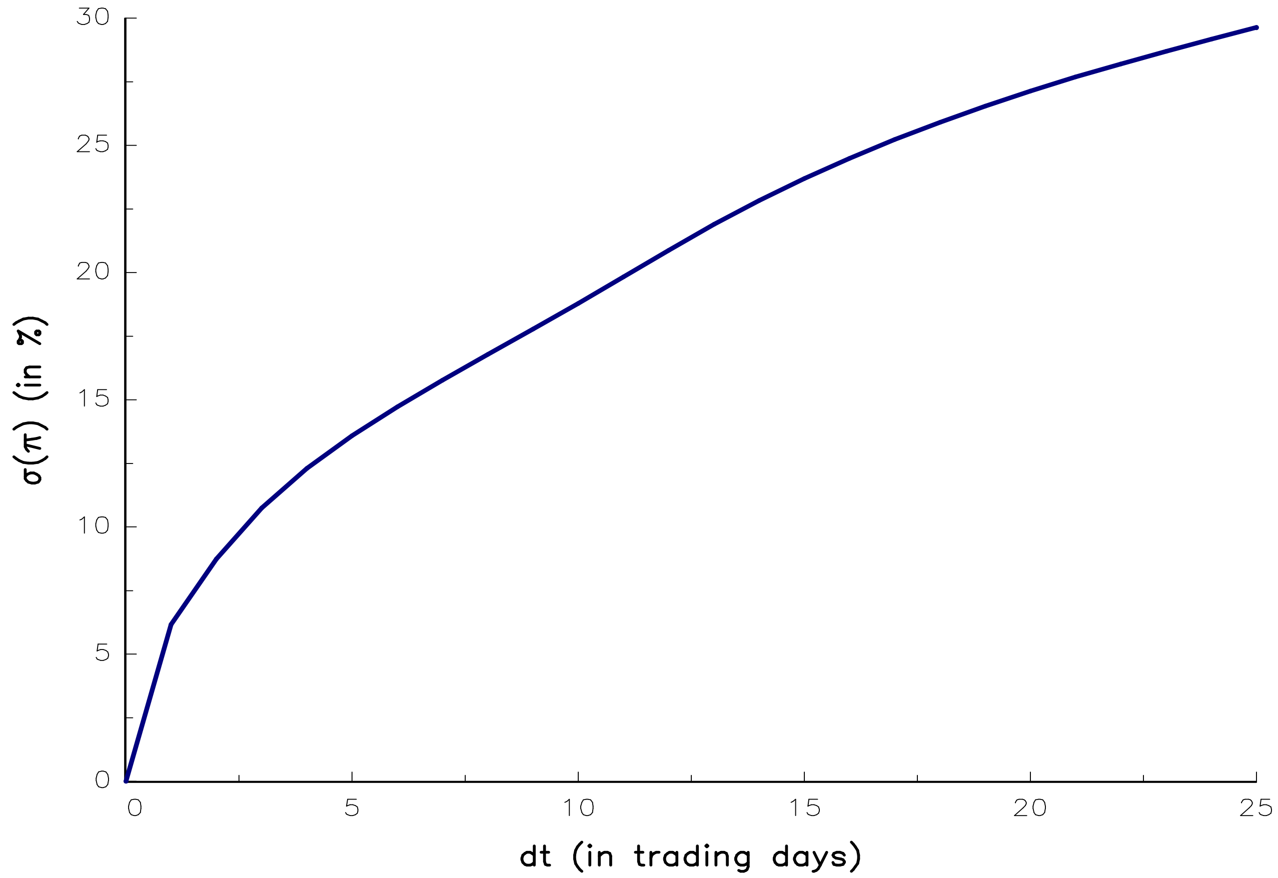


Figure: Relationship between the hedging efficiency $\sigma(\pi)$ and the hedging frequency

Delta hedging

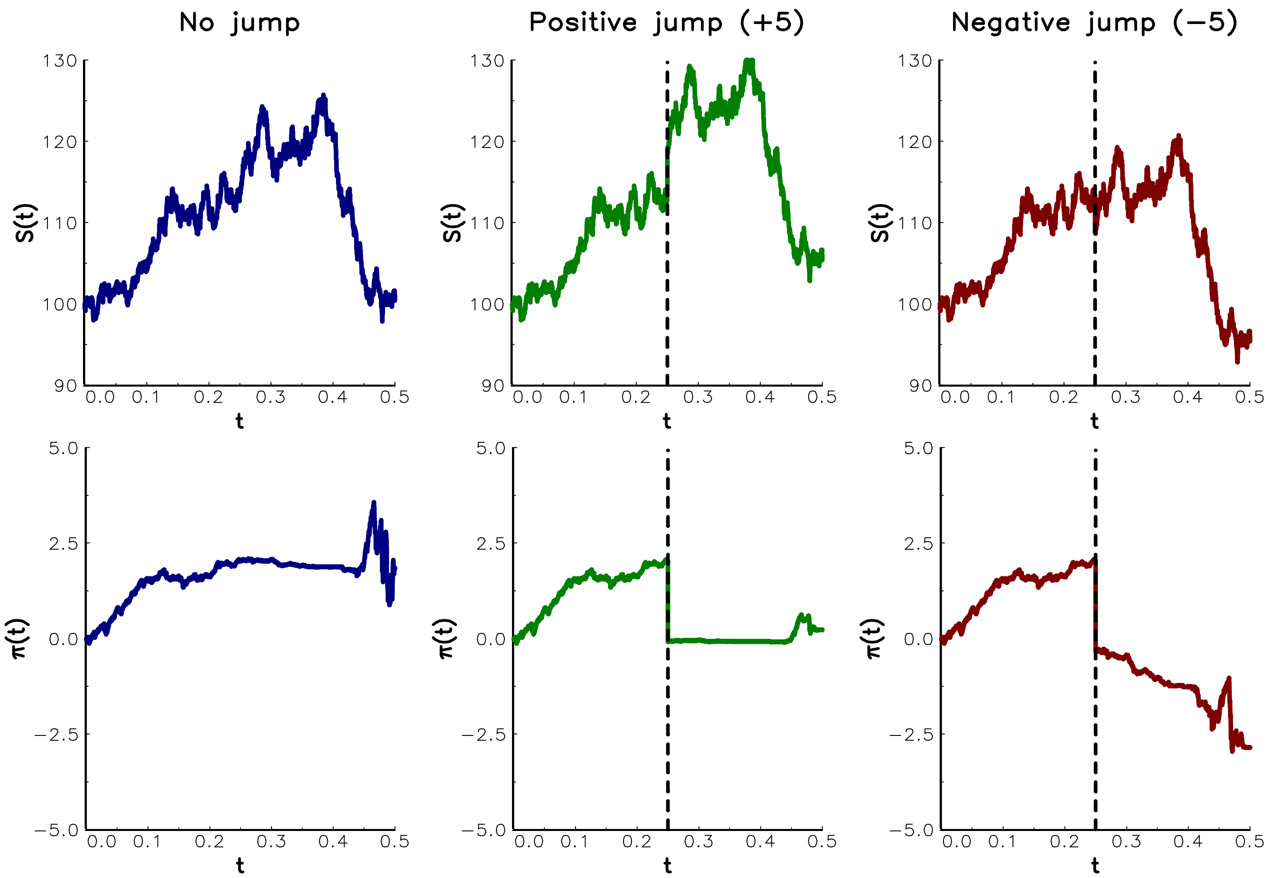


Figure: Impact of a jump on the hedging ratio $\pi(t)$

Delta hedging

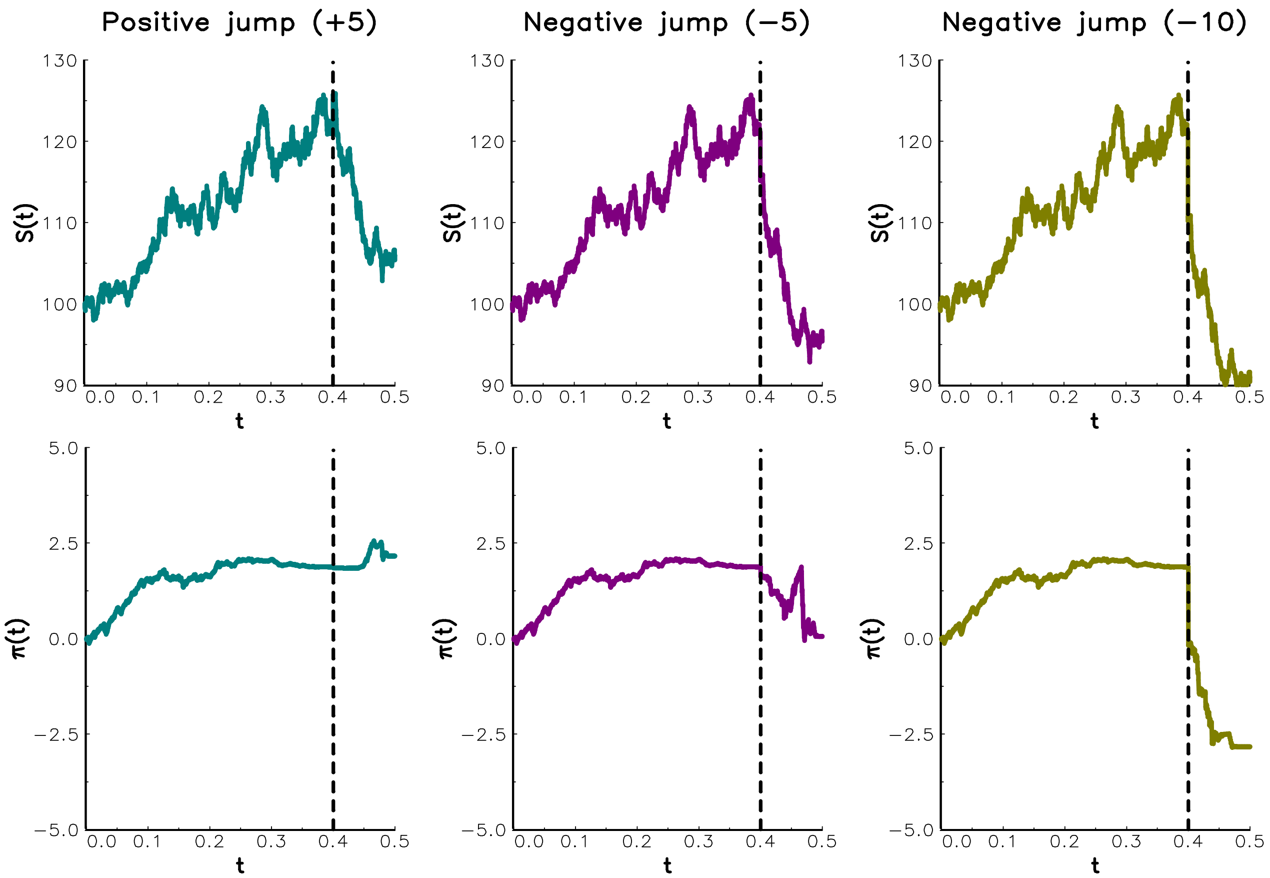


Figure: Impact of a jump on the hedging ratio $\pi(t)$

Greek sensitivities

We have

$$\boldsymbol{\Delta}(t) = \frac{\partial\, \mathcal{C}(t, S(t))}{\partial\, S(t)}$$

and:

$$\mathcal{C}(t + \mathrm{d}t, S(t+h)) - \mathcal{C}(t, S(t)) \approx \boldsymbol{\Delta}(t) \cdot (S(t + \mathrm{d}t) - S(t))$$

$\Rightarrow$ Taylor expansion to other orders and other parameters

Greek sensitivities

The delta-gamma-theta approximation is:

$$\begin{aligned} \mathcal{C}(t+\mathrm{d}t, S(t+h)) - \mathcal{C}(t, S(t)) \approx\ & \boldsymbol{\Delta}(t) \cdot (S(t+\mathrm{d}t) - S(t)) + \\ & \frac{1}{2}\boldsymbol{\Gamma}(t) \cdot (S(t+\mathrm{d}t) - S(t))^2 + \\ & \boldsymbol{\Theta}(t) \cdot ((t+\mathrm{d}t) - t) \end{aligned}$$

where:

$$\boldsymbol{\Gamma}(t) = \frac{\partial^2 \mathcal{C}(t, S(t))}{\partial S(t)^2} = \frac{\partial \boldsymbol{\Delta}(t)}{\partial S(t)}$$

and:

$$\boldsymbol{\Theta}(t) = \frac{\partial \mathcal{C}(t, S(t))}{\partial t} = -\frac{\partial \mathcal{C}(t, S(t))}{\partial T}$$

Greek sensitivities

- We have:

$$\boldsymbol{\Theta}(t) = \frac{\partial\, \mathcal{C}(t, S(t))}{\partial\, t} = -\frac{\partial\, \mathcal{C}(t, S(t))}{\partial\, T}$$

- We recall that the option price satisfies the PDE:

$$\frac{1}{2}\sigma^2 S^2 \boldsymbol{\Gamma} + bS\boldsymbol{\Delta} + \boldsymbol{\Theta} - r\mathcal{C} = 0$$

- We deduce that the theta of the option can be calculated as follows:

$$\boldsymbol{\Theta} = r\mathcal{C} - \frac{1}{2}\sigma^2 S^2 \boldsymbol{\Gamma} - bS\boldsymbol{\Delta}$$

Greek sensitivities

Example #2

We consider a call option, whose strike K is equal to 100. The risk-free rate and the cost-of-carry parameter are equal to 5%. For the volatility coefficient, we consider two cases: (a) $\sigma = 20\%$ and (b) $\sigma = 50\%$.

Greek sensitivities

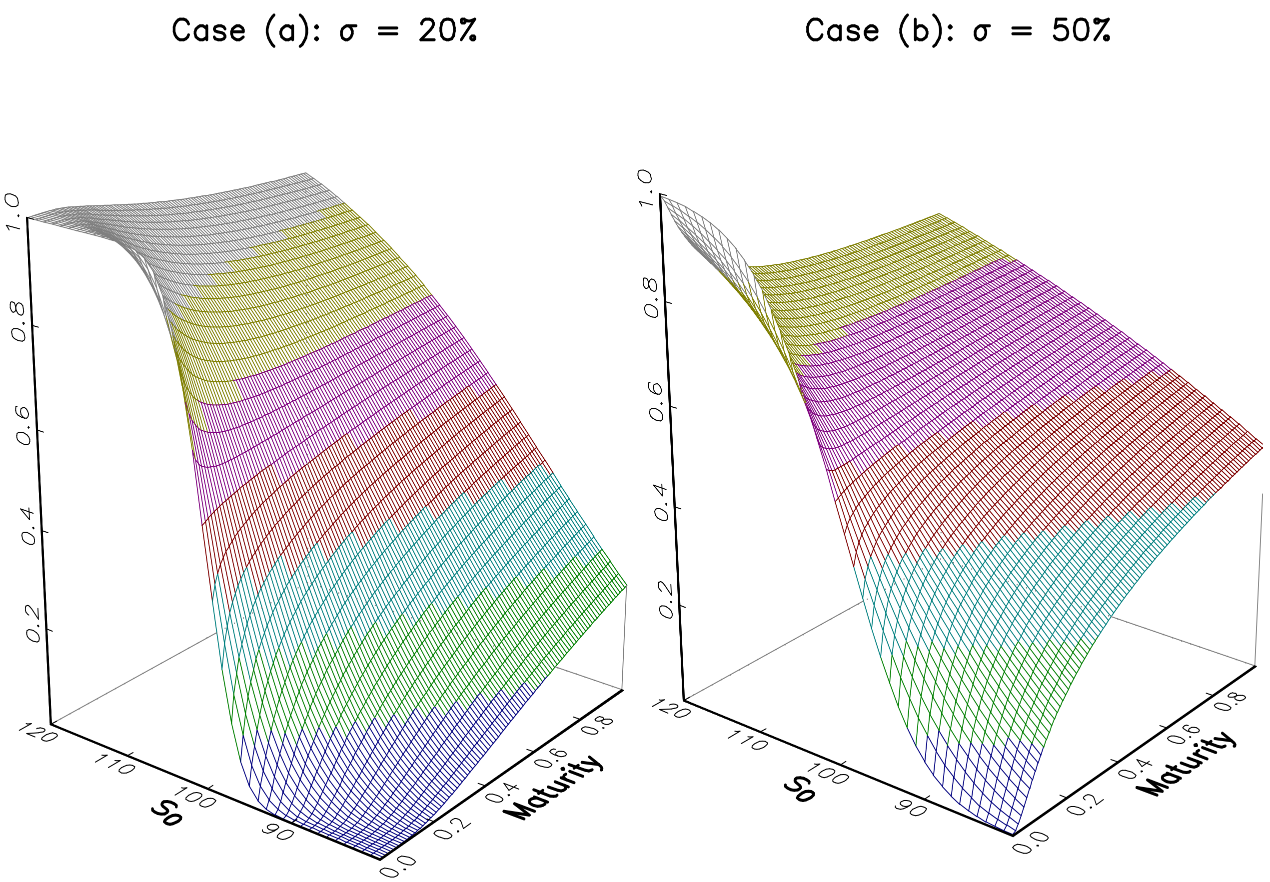


Figure: Delta coefficient of the call option

Greek sensitivities

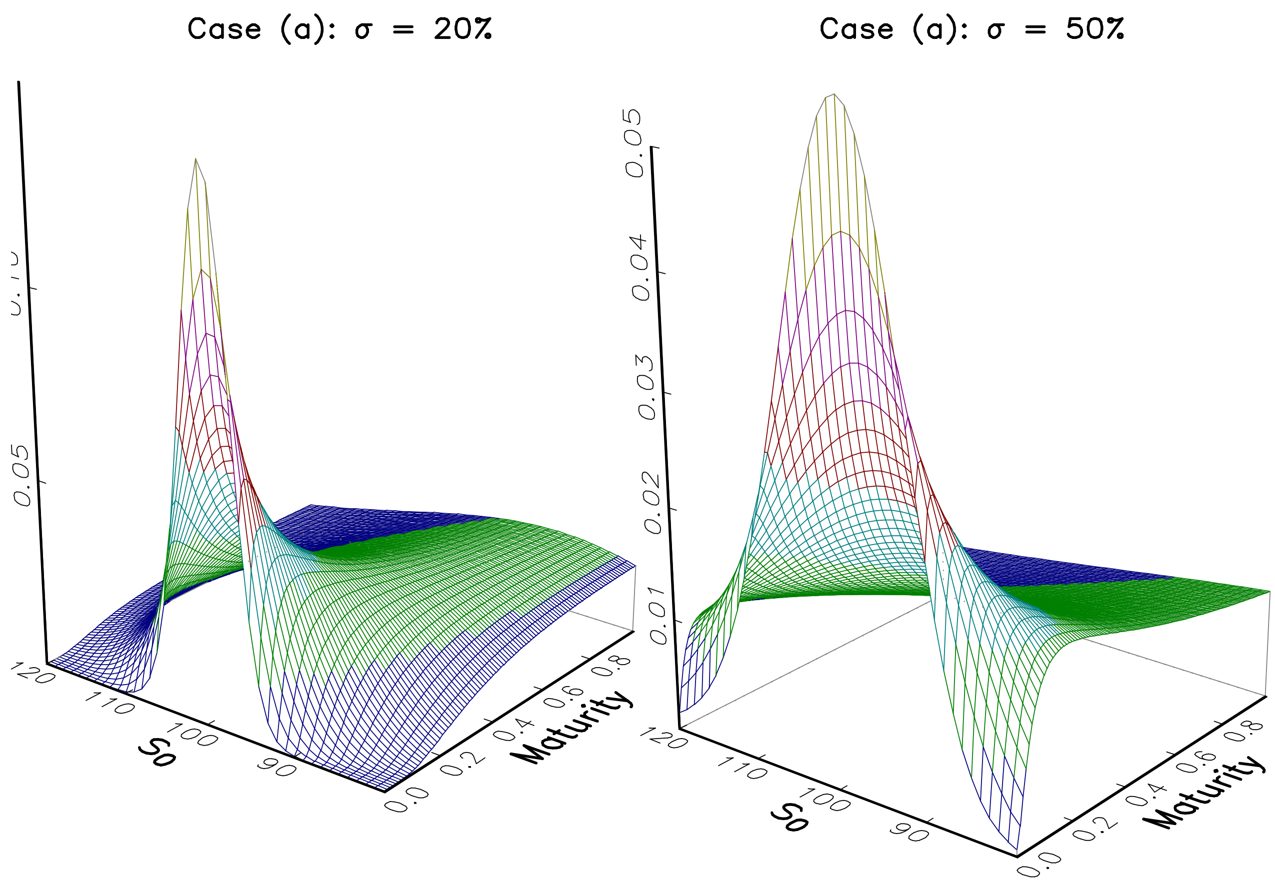


Figure: Gamma coefficient of the call option

Greek sensitivities

- **We assume a delta neutral hedging portfolio**
- The trader can face four configurations of residual risk:

		$\boldsymbol{\Gamma}$ $-$	$\boldsymbol{\Gamma}$ $+$
$\boldsymbol{\Theta}$	$-$		$\checkmark$
$\boldsymbol{\Theta}$	$+$	$\checkmark$	

- The configurations $(\boldsymbol{\Gamma} < 0, \boldsymbol{\Theta} < 0)$ and $(\boldsymbol{\Gamma} > 0, \boldsymbol{\Theta} > 0)$ are not realistic

Greek sensitivities

Two main configurations:

(a) a negative gamma exposure with a positive theta

(b) a positive gamma exposure with a negative theta

Two P&L profiles:

(a) If the gamma is negative, the best situation is obtained when the asset price does not move. Any changes in the asset price reduce the P&L, which can be negative if the gamma effect is more important than the theta effect. We also notice that the gain is bounded and the loss is unbounded in this configuration

(b) If the theta is negative, the loss is bounded and maximum when the asset price does not move. Any changes in the asset price increase the P&L because the gamma is positive. In this configuration, the gain is unbounded

Greek sensitivities

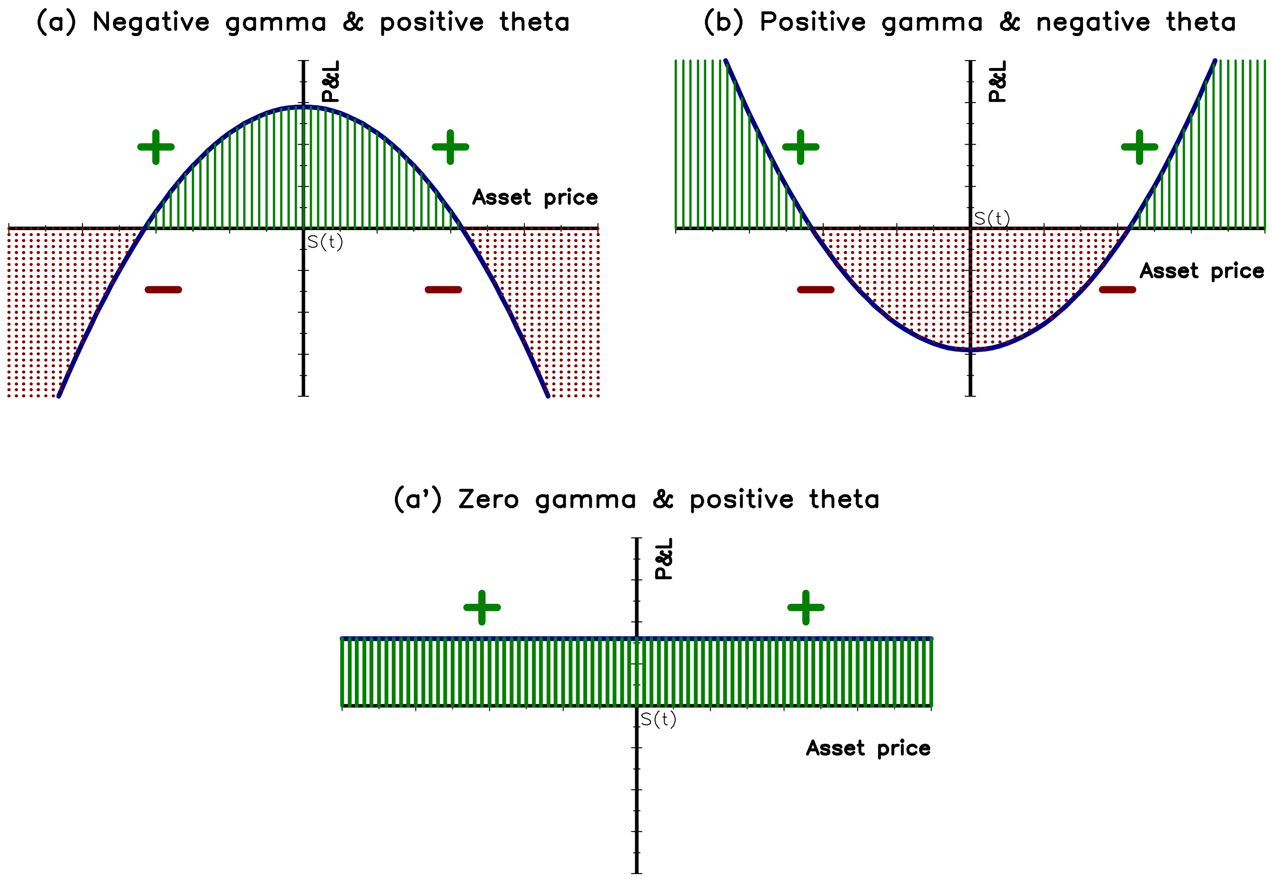


Figure: P&L of the delta neutral hedging portfolio

Greek sensitivities

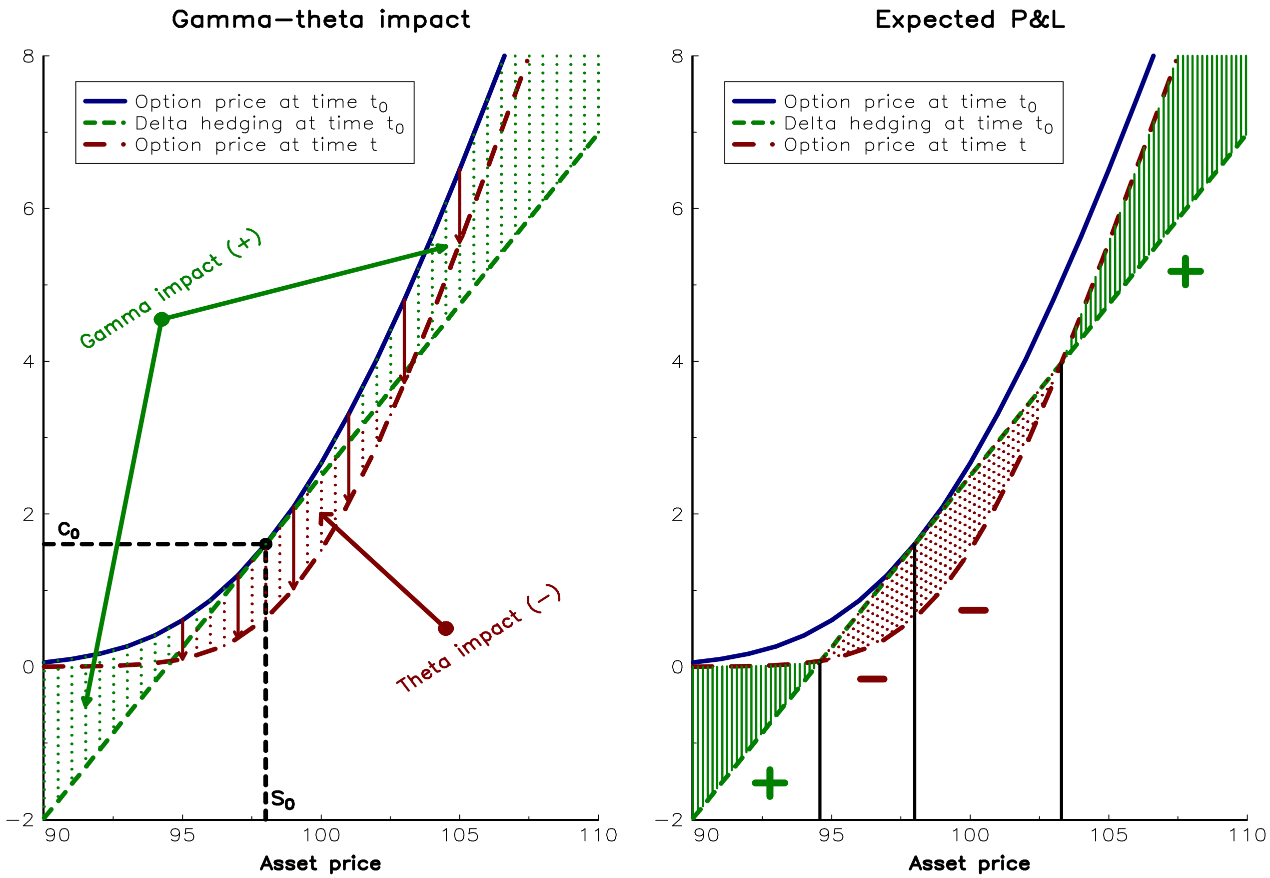


Figure: Illustration of the configuration ($\boldsymbol{\Gamma} > 0, \boldsymbol{\Theta} < 0$)

The implied volatility

Definition

- The implied volatility is the root of the following non-linear equation:

$$f_{\mathrm{BS}}\left(S_0, K, \sigma_{\mathrm{implied}}, T, b, r\right) = V\left(T, K\right)$$

  where f_{BS} is the Black-scholes formula and $V\left(T, K\right)$ is the market price of the option, whose maturity date is T and whose strike is K

- By convention, the implied volatility is denoted by Σ, and is a function of the parameters T and K:

$$\sigma_{\mathrm{implied}} = \Sigma\left(T, K\right)$$

The implied volatility

Example #3

We consider a call option, whose maturity is one year. The current price of the underlying asset is normalized and is equal to 100. Moreover, the risk-free rate and the cost-of-carry parameter are equal to 5%. Below, we report the market price of European call options of three assets for several strikes:

K	90	95	98	100	101	102	105	110
$\mathcal{C}_1(T,K)$	16.70	13.35	11.55	10.45	9.93	9.42	8.02	6.04
$\mathcal{C}_2(T,K)$	18.50	14.50	12.00	10.45	9.60	9.00	7.50	5.70
$\mathcal{C}_3(T,K)$	18.00	14.00	11.80	10.45	9.90	9.50	8.40	7.40

The implied volatility

Table: Implied volatility $\Sigma(T, K)$

K	90	95	98	100	101	102	105	110
$\Sigma_1(T, K)$	20.00	20.01	19.99	20.0	20.01	19.99	20.00	20.00
$\Sigma_2(T, K)$	26.18	23.41	21.24	20.0	19.14	18.90	18.69	19.14
$\Sigma_3(T, K)$	24.53	21.95	20.68	20.0	19.93	20.20	20.95	23.43

The implied volatility

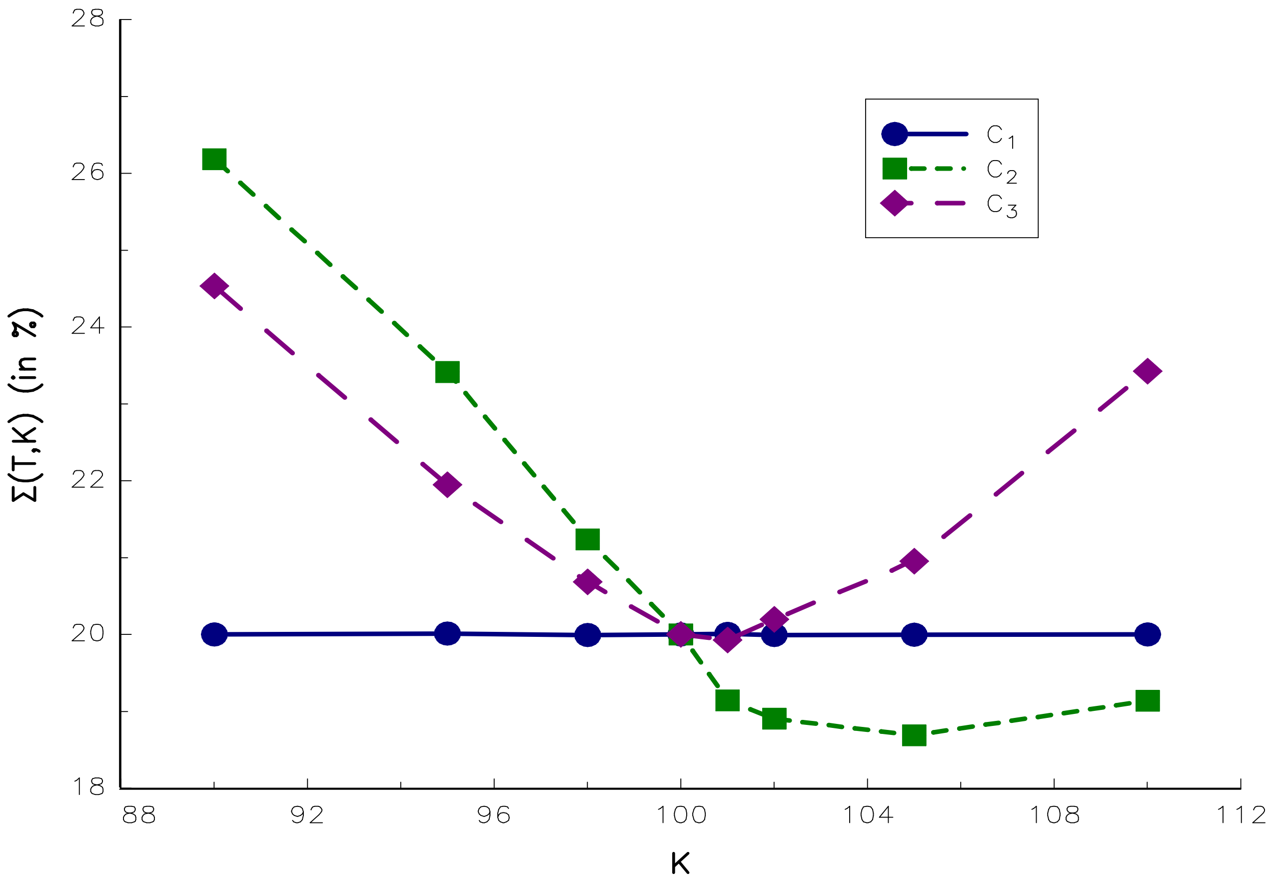


Figure: Volatility smile

The implied volatility

- When the curve of implied volatility is decreasing and increasing, the curve is called a volatility smile
- When the curve of implied volatility is just decreasing, it is called a volatility skew
- If we consider the maturity dimension, the term structure of implied volatility is known as the volatility surface

Relationship between the implied volatility and the risk-neutral density

- We have:

$$\begin{aligned}
\mathcal{C}_t(T,K) &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r\,\mathrm{d}s} \left(S(T)-K\right)^+ \middle| \mathcal{F}_t \right] \\
&= e^{-r(T-t)} \int_{-\infty}^{\infty} (S-K)^+ q_t(T,S)\,\mathrm{d}S \\
&= e^{-r(T-t)} \int_{K}^{\infty} (S-K)\, q_t(T,S)\,\mathrm{d}S
\end{aligned}$$

  where $q_t(T,S)$ is the risk-neutral probability density function of $S(T)$ at time t
- By definition, the risk-neutral cumulative distribution function $\mathbb{Q}_t(T,S)$ is equal to:

$$\mathbb{Q}_t(T,S) = \int_{-\infty}^{S} q_t(T,x)\,\mathrm{d}x$$

Relationship between the implied volatility and the risk-neutral density

- We deduce that:

$$\begin{aligned}\frac{\partial\,\mathcal{C}_t\left(T,K\right)}{\partial\,K} &= -e^{-r(T-t)}\int_K^{\infty} q_t\left(T,S\right)\,\mathrm{d}S \\ &= -e^{-r(T-t)}\left(1-\mathbb{Q}_t\left(T,K\right)\right)\end{aligned}$$

and:

$$\frac{\partial^2\,\mathcal{C}_t\left(T,K\right)}{\partial\,K^2} = e^{-r(T-t)}q_t\left(T,K\right)$$

- It follows that:

$$\begin{aligned}\mathbb{Q}_t\left(T,K\right) &= \Pr\left\{S\left(T\right)\leq K \mid \mathcal{F}_t\right\} \\ &= 1+e^{r(T-t)}\cdot\partial_K\mathcal{C}_t\left(T,K\right)\end{aligned}$$

Relationship between the implied volatility and the risk-neutral density

- We note $\Sigma_t(T, K)$ the volatility surface and $\mathcal{C}_t^\star(T, K, \Sigma)$ the Black-Scholes formula. It follows that:

$$\begin{aligned}\mathbb{Q}_t(T, K) = & \; 1 + e^{r(T-t)} \cdot \partial_K \mathcal{C}_t^\star(T, K, \Sigma_t(T, K)) + \\ & e^{r(T-t)} \cdot \partial_\Sigma \mathcal{C}_t^\star(T, K, \Sigma_t(T, K)) \cdot \partial_K \Sigma_t(T, K)\end{aligned}$$

where:

$$\partial_K \mathcal{C}_t^\star(T, K, \Sigma) = -e^{-r(T-t)} \cdot \Phi(d_2)$$

and:

$$\partial_\Sigma \mathcal{C}_t^\star(T, K, \Sigma) = S(t) \cdot e^{(b-r)(T-t)} \cdot \sqrt{T-t} \cdot \phi\left(d_2 + \Sigma\sqrt{T-t}\right)$$

Relationship between the implied volatility and the risk-neutral density

- The risk-neutral probability density function is equal to:

$$q_t(T,K) = \partial_K \mathbb{Q}_t(T,K) = e^{r(T-t)} \cdot \partial_K^2 \mathcal{C}_t(T,K)$$

where:

$$\begin{aligned}\partial_K^2 \mathcal{C}_t(T,K) &= \partial_K^2 \mathcal{C}_t^\star(T,K,\Sigma_t) + 2\cdot\partial_{K,\Sigma}^2 \mathcal{C}_t^\star(T,K,\Sigma_t)\cdot\partial_K\Sigma_t(T,K) + \\ & \quad \partial_\Sigma \mathcal{C}_t^\star(T,K,\Sigma_t)\cdot\partial_K^2\Sigma_t(T,K) + \\ & \quad \partial_\Sigma^2 \mathcal{C}_t^\star(T,K,\Sigma_t)\cdot(\partial_K\Sigma_t(T,K))^2\end{aligned}$$

and:

$$\begin{aligned}\partial_K^2 \mathcal{C}_t^\star(T,K,\Sigma) &= e^{-r(T-t)}\frac{\phi(d_2)}{K\Sigma\sqrt{T-t}} \\ \partial_{K,\Sigma}^2 \mathcal{C}_t^\star(T,K,\Sigma) &= e^{(b-r)(T-t)}\frac{S(t)\,d_1\phi(d_1)}{\Sigma K} \\ \partial_\Sigma^2 \mathcal{C}_t^\star(T,K,\Sigma) &= e^{(b-r)(T-t)}\frac{S(t)\,d_1 d_2\sqrt{T-t}\phi(d_1)}{\Sigma}\end{aligned}$$

Relationship between the implied volatility and the risk-neutral density

Example #4

We assume that $S(t) = 100$, $T - t = 10$, $b = r = 5\%$ and:

$$\Sigma_t(T, K) = 0.25 + \ln\left(1 + 10^{-6}(K - 90)^2 + 10^{-6}(K - 180)^2\right)$$

Relationship between the implied volatility and the risk-neutral density

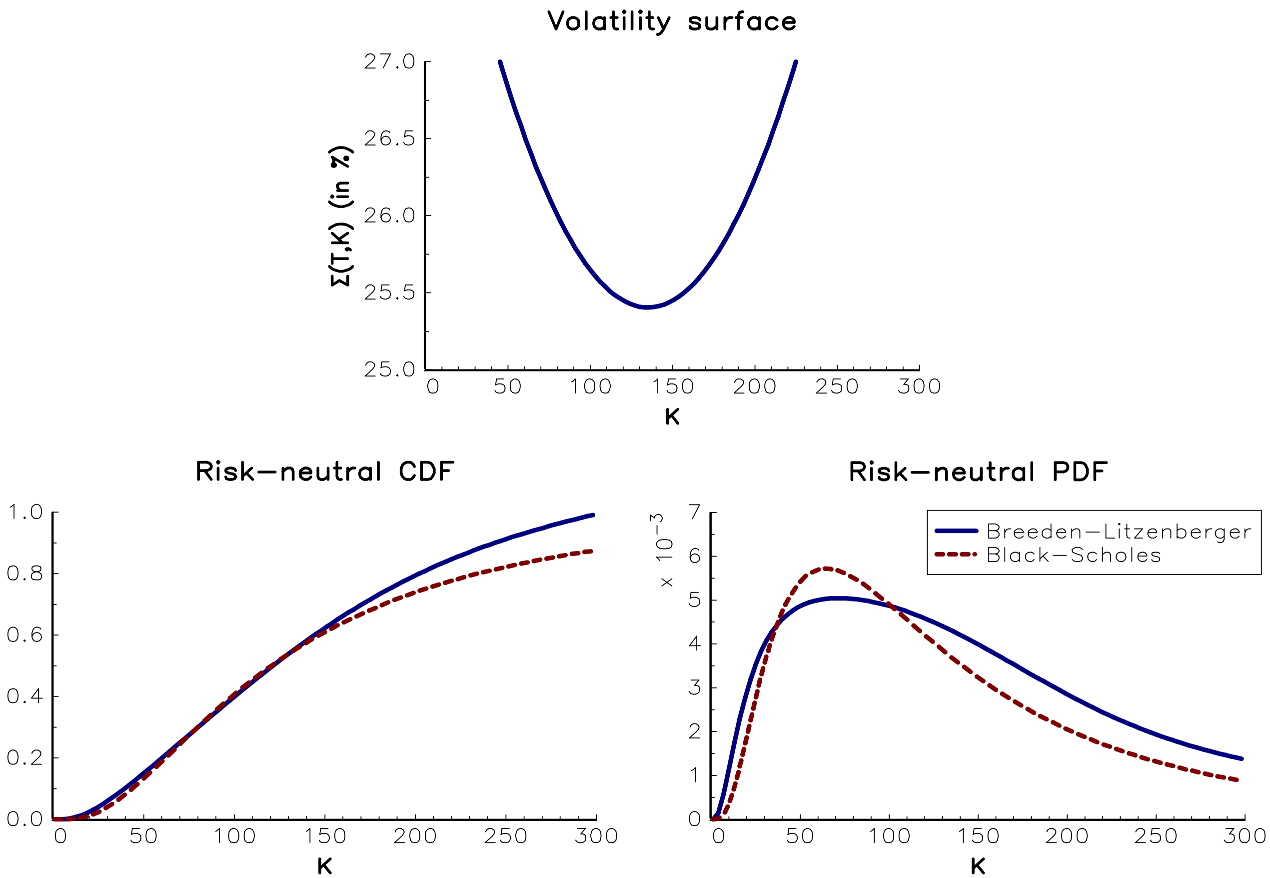


Figure: Risk-neutral probability density function

Robustness of the Black-Scholes formula

We can show that:

$$V(T) - f(S(T)) = \frac{1}{2}\int_0^T e^{r(T-t)} \boldsymbol{\Gamma}(t)\left(\Sigma^2(T,K) - \sigma^2(t)\right) S^2(t)\,\mathrm{d}t$$

where $f(S(T))$ is the payoff of the option. We obtain the following results:

- if $\boldsymbol{\Gamma}(t) \geq 0$, a positive P&L is achieved by overestimating the realized volatility:
$$\Sigma(T,K) \geq \sigma(t) \Longrightarrow V(T) \geq f(S(T))$$
- if $\boldsymbol{\Gamma}(t) \leq 0$, a positive P&L is achieved by underestimating the realized volatility:
$$\Sigma(T,K) \leq \sigma(t) \Longrightarrow V(T) \geq f(S(T))$$
- the variance of the hedging error is an increasing function of the absolute value of the gamma coefficient:
$$|\boldsymbol{\Gamma}(t)| \nearrow \Longrightarrow \operatorname{var}(V(T) - f(S(T))) \nearrow$$

Robustness of the Black-Scholes formula

Example #5

We consider the replication of 100 ATM call options. The current price of the asset is 100 and the maturity of the option is 6 months (or 130 trading days). We consider the following parameters: $b = r = 5\%$. We rebalance the delta hedging portfolio every trading day. Moreover, we assume that the option is priced and hedged with a 20% implied volatility.

Relationship between the implied volatility and the risk-neutral density

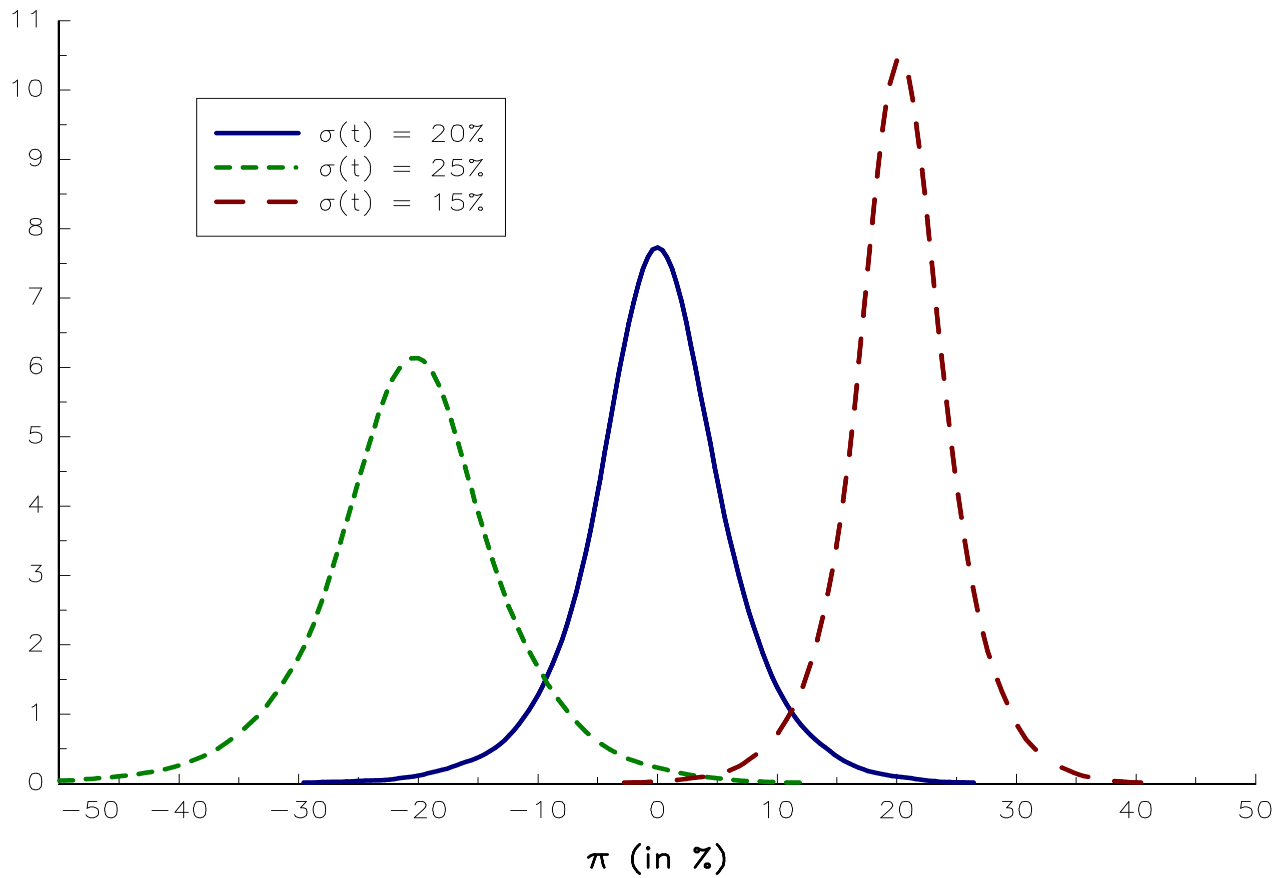


Figure: Hedging error when the implied volatility is 20%

Robustness of the Black-Scholes formula

We obtain:

$$\Pr\{\pi > 0 \mid \Sigma = 20\%, \sigma = 15\%\} = 99.04\%$$

and:

$$\Pr\{\pi > 0 \mid \Sigma = 20\%, \sigma = 25\%\} = 0.09\%$$

Vasicek model

- The instantaneous interest rate follows an Ornstein-Uhlenbeck process:

$$\begin{cases} \mathrm{d}r(t) = a(b - r(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W(t) \\ r(t_0) = r_0 \end{cases}$$

- The value $V(t,r)$ of a zero-coupon bond satisfies the PDE:

$$\frac{1}{2}\sigma^2\frac{\partial^2 V(t,r)}{\partial r^2} + (a(b - r(t)) - \lambda(t)\sigma)\frac{\partial V(t,r)}{\partial r} + \frac{\partial V(t,r)}{\partial t} - r(t)V(t,r) = 0$$

with $V(T,r) = 1$

- The Feynman-Kac formula implies:

$$V(0,r_0) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r(t)\,\mathrm{d}t} \,\middle|\, \mathcal{F}_0 \right]$$

where $\mathrm{d}r(t) = (a(b - r(t)) - \lambda(t)\sigma)\,\mathrm{d}t + \sigma\,\mathrm{d}W^{\mathbb{Q}}(t)$

Vasicek model

By assuming that $\lambda(t) = \lambda$, we have:

$$V(0, r_0) = \exp\left(-r_0\beta - \left(b' - \frac{\sigma^2}{2a^2}\right)(T - \beta) - \frac{\sigma^2\beta^2}{4a}\right)$$

where $b' = b - \dfrac{\lambda\sigma}{a}$ and $\beta = \dfrac{1 - e^{-aT}}{a}$

Vasicek model

We recall that the zero-coupon rate is defined by:

$$B(t,T) = e^{-(T-t)R(t,T)}$$

We deduce that:

$$\begin{aligned} R(t,T) &= -\frac{1}{T-t}\ln B(t,T) \\ &= \left(b' - \frac{\sigma^2}{2a^2}\right) + \left(r_t - b' + \frac{\sigma^2}{2a^2}\right)\frac{\beta}{T-t} + \frac{\sigma^2\beta^2}{4a(T-t)} \end{aligned}$$

Vasicek model

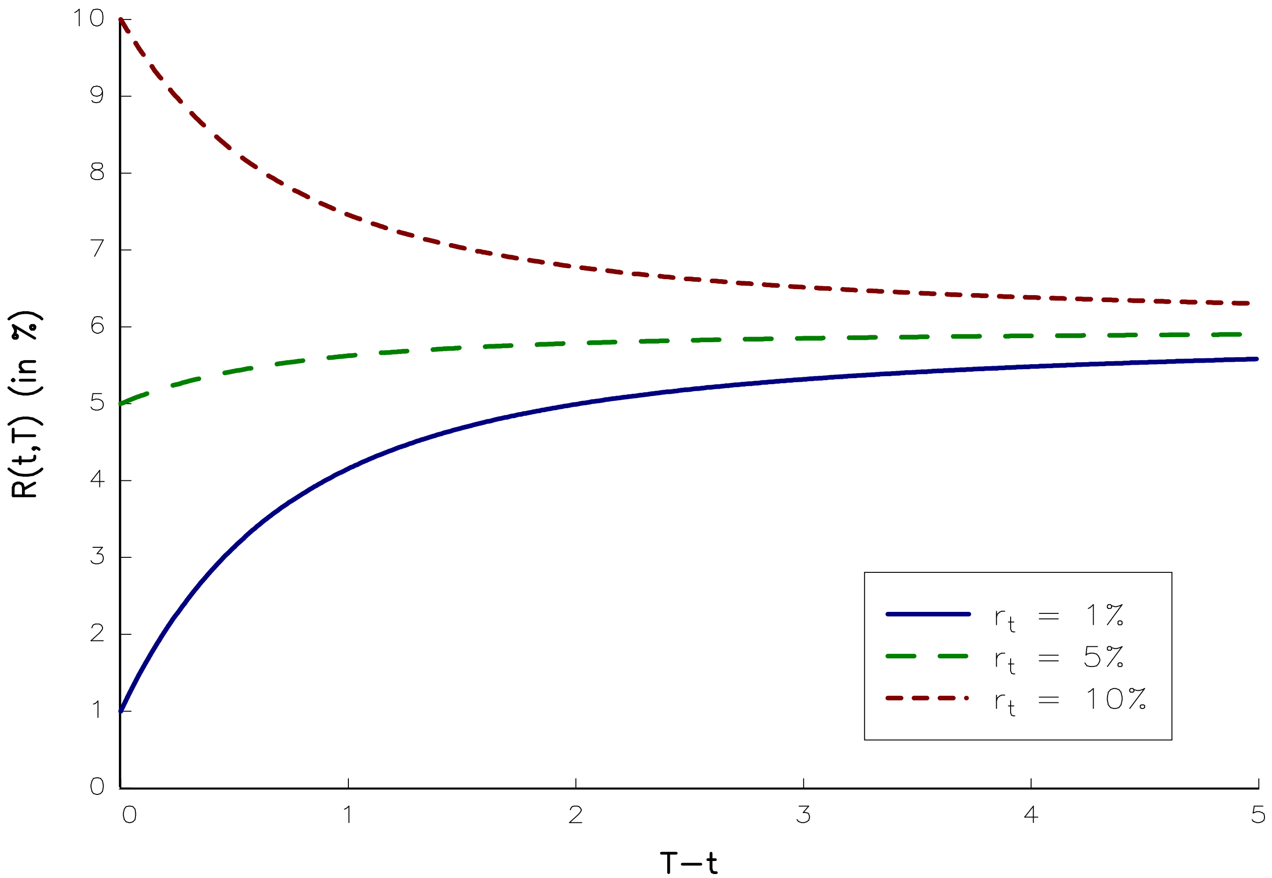


Figure: Vasicek model ($a = 2.5$, $b = 6\%$ and $\sigma = 5\%$)

Vasicek model

- Let $F(t, T_1, T_2)$ be the forward rate at time t for the period $[T_1, T_2]$:

$$B(t, T_2) = e^{-(T_2 - T_1)F(t, T_1, T_2)} B(t, T_1)$$

- We deduce that the expression of $F(t, T_1, T_2)$ is:

$$F(t, T_1, T_2) = -\frac{1}{(T_2 - T_1)} \ln \frac{B(t, T_2)}{B(t, T_1)}$$

- It follows that the instantaneous forward rate is given by this equation:

$$f(t, T) = F(t, T, T) = -\frac{\partial \ln B(t, T)}{\partial T}$$

Vasicek model

Another expression of the price of the zero-coupon bond is:

$$B(t, r_t) = \exp\left(-(T-t)R_\infty - (r_t - R_\infty)\left(\frac{1-e^{-a(T-t)}}{a}\right) - \frac{\sigma^2\left(1-e^{-a(T-t)}\right)^2}{4a^3}\right)$$

where:

$$R_\infty = \lim_{T\to\infty} R(t, T) = b' - \frac{\sigma^2}{2a^2}$$

Therefore, the instantaneous forward rate in the Vasicek model is:

$$f(t, T) = R_\infty + (r_t - R_\infty)e^{-a(T-t)} + \frac{\sigma^2\left(1-e^{-a(T-t)}\right)e^{-a(T-t)}}{2a^2}$$

Vasicek model

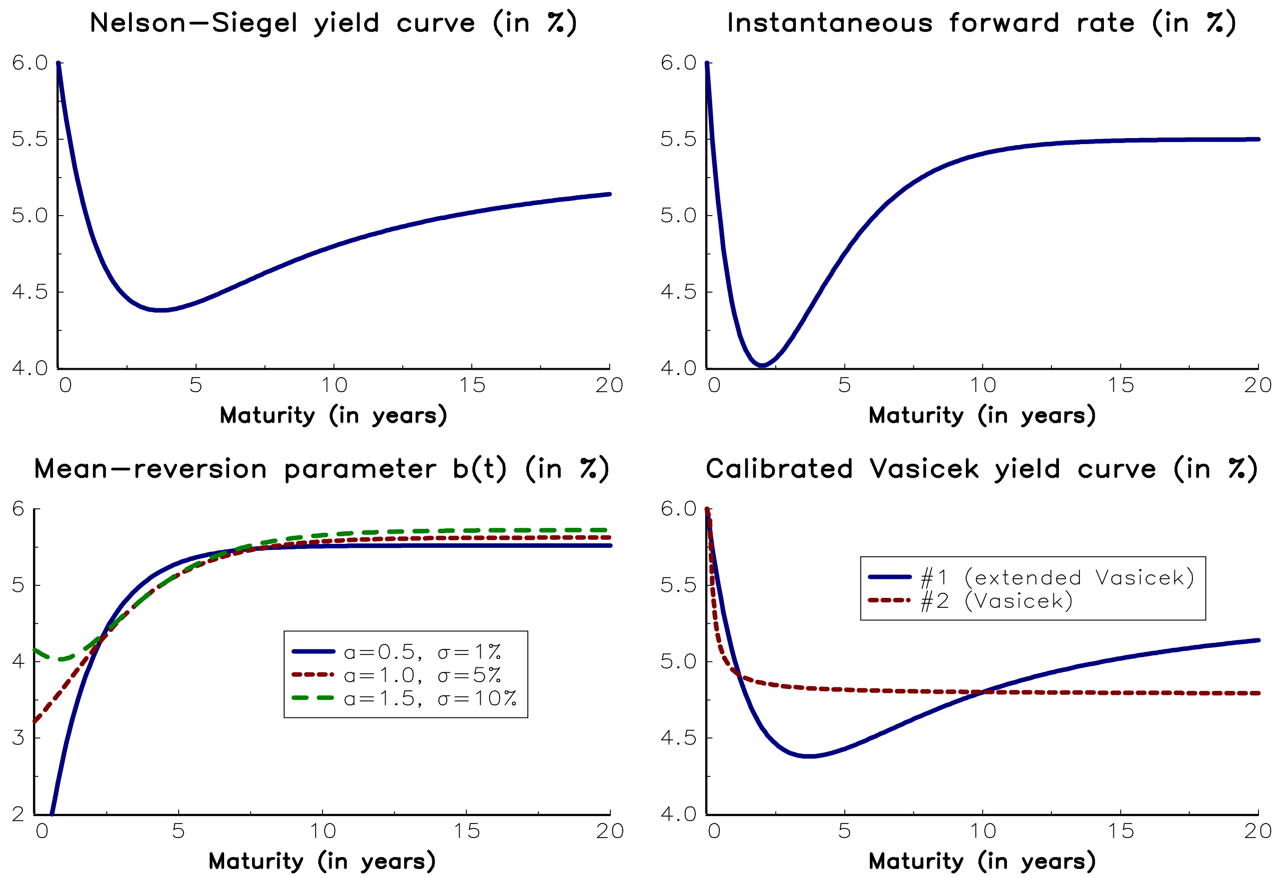


Figure: Calibration of the Vasicek model

Caps, floors and swaptions

- Future dates $T_0, T_1, \ldots, T_n$
- The payoff of a caplet is $(T_i - T_{i-1}) \left(F(T_{i-1}, T_{i-1}, T_i) - K \right)^+$, where K is the strike of the caplet and $F(T_{i-1}, T_{i-1}, T_i)$ is the forward rate at the future date T_{i-1}
- $\delta_{i-1} = T_i - T_{i-1}$ is the tenor of the caplet
- T_{i-1} is the resetting date (or the fixing date) of the forward rate
- T_i is the maturity date of the caplet
- A cap is a portfolio of successive caplets:

$$\mathrm{Cap}(t) = \sum_{i=1}^{n} \mathrm{Caplet}(t, T_{i-1}, T_i)$$

Caps, floors and swaptions

A floor is a portfolio of successive floorlets:

$$\text{Floor}(t) = \sum_{i=1}^{n} \text{Floorlet}(t, T_{i-1}, T_i)$$

where the payoff of the floorlet is $(T_i - T_{i-1})(K - F(T_{i-1}, T_{i-1}, T_i))^+$

Caps, floors and swaptions

A par swap rate is the fixed rate of an interest rate swap:

$$\mathrm{Sw}(t) = \frac{B(t, T_0) - B(t, T_n)}{\sum_{i=1}^{n} (T_i - T_{i-1}) \cdot B(t, T_i)}$$

The payoff of a payer swaption is:

$$(\mathrm{Sw}(T_0) - K)^{+} \sum_{i=1}^{n} (T_i - T_{i-1}) B(T_0, T_i)$$

where $\mathrm{Sw}(T_0)$ is the forward swap rate

Caps, floors and swaptions

Generally, caps, floors and swaptions are written on the Libor rate, which is defined as a simple forward rate:

$$L(t, T_{i-1}, T_i) = \frac{1}{T_i - T_{i-1}} \left(\frac{B(t, T_{i-1})}{B(t, T_i)} - 1 \right)$$

We have:

$$\mathrm{Caplet}(t, T_{i-1}, T_i) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_i} r(s)\,\mathrm{d}s} \delta_{i-1} \left(L(T_{i-1}, T_{i-1}, T_i) - K \right)^+ \middle| \mathcal{F}_t \right]$$

and:

$$\mathrm{Swaption}(t) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_n} r(s)\,\mathrm{d}s} \left(\mathrm{Sw}(T_0) - K \right)^+ \sum_{i=1}^{n} \delta_{i-1} B(T_0, T_i) \middle| \mathcal{F}_t \right]$$

$\Rightarrow$ **the discount factor is stochastic** and is not independent from the forward rate $L(T_{i-1}, T_{i-1}, T_i)$ or the forward swap rate $\mathrm{Sw}(T_0)$

Change of numéraire and equivalent martingale measure

- The price of the contingent claim, whose payoff is $V(T) = f(S(T))$ at time T, is given by:

$$V(0) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r(s)\,\mathrm{d}s} \cdot V(T) \,\middle|\, \mathcal{F}_0 \right]$$

- $\mathbb{Q}$ is the risk-neutral probability measure
- We can rewrite this equation as follows:

$$\frac{V(0)}{M(0)} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V(T)}{M(T)} \,\middle|\, \mathcal{F}_0 \right]$$

where $M(0) = 1$ and:

$$M(t) = \exp\left(\int_0^t r(s)\,\mathrm{d}s \right)$$

- Under the probability measure $\mathbb{Q}$, we know that $\tilde{V}(t) = V(t)/M(t)$ is an $\mathcal{F}_t$-martingale

Change of numéraire and equivalent martingale measure

- The money market account $M(t)$ is then the numéraire when the martingale measure is the risk-neutral probability measure, but other numéraires can be used in order to simplify pricing problems:

  "The use of the risk-neutral probability measure has proved to be very powerful for computing the prices of contingent claims [...] We show here that many other probability measures can be defined in the same way to solve different asset-pricing problems, in particular option pricing. Moreover, these probability measure changes are in fact associated with numéraire changes" (Geman et al., 1995, page 443).

Change of numéraire and equivalent martingale measure

Let us consider another numéraire $N(t) > 0$ and the associated probability measure given by the Radon-Nikodym derivative:

$$\frac{\mathrm{d}\mathbb{Q}^{\star}}{\mathrm{d}\mathbb{Q}} = \frac{N(T)/N(0)}{M(T)/M(0)} = e^{-\int_0^T r(s)\,\mathrm{d}s} \cdot \frac{N(T)}{N(0)}$$

We have:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{\star}}\left[\left.\frac{V(T)}{N(T)}\right| \mathcal{F}_0\right] &= \mathbb{E}^{\mathbb{Q}}\left[\left.\frac{V(T)}{N(T)} \cdot \frac{\mathrm{d}\mathbb{Q}^{\star}}{\mathrm{d}\mathbb{Q}}\right| \mathcal{F}_t\right] \\
&= \frac{M(0)}{N(0)} \cdot \mathbb{E}^{\mathbb{Q}}\left[\left.\frac{V(T)}{M(T)}\right| \mathcal{F}_0\right] \\
&= \frac{M(0)}{N(0)} \cdot V(0)
\end{aligned}$$

Change of numéraire and equivalent martingale measure

- We deduce that:

$$\frac{V(0)}{N(0)} = \mathbb{E}^{\mathbb{Q}^\star}\left[\left.\frac{V(T)}{N(T)}\right| \mathcal{F}_0\right]$$

- We have changed the numéraire $(M(t) \rightarrow N(t))$ and the probability measure $(\mathbb{Q} \rightarrow \mathbb{Q}^\star)$
- More generally, we have:

$$V(t) = N(t) \cdot \mathbb{E}^{\mathbb{Q}^\star}\left[\left.\frac{V(T)}{N(T)}\right| \mathcal{F}_t\right]$$

- Thanks to Girsanov theorem, we notice that $e^{-\int_0^t r(s)\,\mathrm{d}s} N(t)$ is an $\mathcal{F}_t$-martingale

Change of numéraire and equivalent martingale measure

- The forward numéraire is the zero-coupon bond price of maturity T:

$$N(t) = B(t, T)$$

- The probability measure is called the forward probability and is denoted by $\mathbb{Q}^{\star}(T)$
- By noticing that $N(T) = B(T, T) = 1$, we have:

$$V(t) = B(t, T)\,\mathbb{E}^{\mathbb{Q}^{\star}(T)}\left[V(T) \mid \mathcal{F}_t\right]$$

Change of numéraire and equivalent martingale measure

- In the case of a caplet, we obtain:

$$\begin{aligned}\text{Caplet}(t, T_{i-1}, T_i) &= \delta_{i-1}\mathbb{E}^{\mathbb{Q}}\left[\frac{M(t)}{M(T_i)}(L(T_{i-1}, T_{i-1}, T_i) - K)^+ \,\middle|\, \mathcal{F}_t\right] \\ &= \delta_{i-1}\mathbb{E}^{\mathbb{Q}^\star(T_i)}\left[\frac{N(t)}{N(T_i)}(L(T_{i-1}, T_{i-1}, T_i) - K)^+ \,\middle|\, \mathcal{F}_t\right] \\ &= \delta_{i-1}B(t, T_i)\,\mathbb{E}^{\mathbb{Q}^\star(T_i)}\left[(L(T_{i-1}, T_{i-1}, T_i) - K)^+ \,\middle|\, \mathcal{F}_t\right]\end{aligned}$$

  where $L(t, T_{i-1}, T_i)$ is an $\mathcal{F}_t$-martingale under the forward probability measure $\mathbb{Q}^\star(T_i)$

- The general formula of the caplet price is:

$$\text{Caplet}(t, T_{i-1}, T_i) = B(t, T_i)\,\mathbb{E}^{\mathbb{Q}^\star(T_i)}\left[\left(\frac{1}{B(T_{i-1}, T_i)} - (1 + \delta_{i-1}K)\right)^+ \,\middle|\, \mathcal{F}_t\right]$$

Change of numéraire and equivalent martingale measure

If we use the standard Black model, we obtain:

$$\mathrm{Caplet}\left(t, T_{i-1}, T_i\right) = \delta_{i-1} B\left(t, T_i\right)\left(L\left(t, T_{i-1}, T_i\right)\Phi\left(d_1\right) - K\Phi\left(d_2\right)\right)$$

where σ_{i-1} is the volatility of the Libor rate $L\left(t, T_{i-1}, T_i\right)$,

$$d_1 = \frac{1}{\sigma_{i-1}\sqrt{T_{i-1} - t}} \ln \frac{L\left(t, T_{i-1}, T_i\right)}{K} + \frac{1}{2}\sigma_{i-1}\sqrt{T_{i-1} - t}$$

and:

$$d_2 = d_1 - \sigma_{i-1}\sqrt{T_{i-1} - t}$$

Change of numéraire and equivalent martingale measure

- The annuity numéraire is equal to:

$$N(t) = \sum_{i=1}^{n} (T_i - T_{i-1}) B(t, T_i)$$

- While the forward swap rate is a martingale under the annuity probability measure $\mathbb{Q}^\star$, the price of the swaption is:

$$\begin{aligned}
\mathrm{Swaption}(t) &= \mathbb{E}^{\mathbb{Q}}\left[\left.\frac{M(t)}{M(T_n)}(\mathrm{Sw}(T_0) - K)^+ \sum_{i=1}^{n} \delta_{i-1} B(T_0, T_i) \right| \mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}^\star}\left[\left.\frac{N(t)}{N(T_0)}(\mathrm{Sw}(T_0) - K)^+ \sum_{i=1}^{n} \delta_{i-1} B(T_0, T_i) \right| \mathcal{F}_t\right] \\
&= N(t)\, \mathbb{E}^{\mathbb{Q}^\star}\left[\left.(\mathrm{Sw}(T_0) - K)^+ \right| \mathcal{F}_t\right] \\
&= N(t)\, \mathbb{E}^{\mathbb{Q}^\star}\left[\left.\left(\frac{1 - B(T_0, T_n)}{N(T_0)} - K\right)^+ \right| \mathcal{F}_t\right]
\end{aligned}$$

The HJM model

- Under the risk-neutral probability measure $\mathbb{Q}$, the dynamics of the instantaneous forward rate for the maturity T is given by:

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)\,\mathrm{d}s + \int_0^t \sigma(s,T)\,\mathrm{d}W^{\mathbb{Q}}(s)$$

  where $f(0,T)$ is the current forward rate

- Therefore, the stochastic differential equation is:

$$\mathrm{d}f(t,T) = \alpha(t,T)\,\mathrm{d}t + \sigma(t,T)\,\mathrm{d}W^{\mathbb{Q}}(t)$$

The HJM model

- We can show that:

$$\alpha(t, T) = \sigma(t, T) \int_t^T \sigma(t, u) \, \mathrm{d}u$$

This equation is known as the '*drift restriction*' and is necessary to ensure no-arbitrage opportunities

- We have:

$$\mathrm{d}B(t, T) = r(t) B(t, T) \, \mathrm{d}t + b(t, T) B(t, T) \, \mathrm{d}W^{\mathbb{Q}}(t)$$

where $b(t, T) = -\int_t^T \sigma(t, u) \, \mathrm{d}u$

The HJM model

- The drift restriction implies that the dynamics of the instantaneous forward rate $f(t,T)$ is given by:

$$\mathrm{d}f(t,T) = \left(\sigma(t,T)\int_t^T \sigma(t,u)\,\mathrm{d}u\right)\mathrm{d}t + \sigma(t,T)\,\mathrm{d}W^{\mathbb{Q}}(t)$$

- If we are interested in the instantaneous spot rate $r(t)$, we obtain:

$$\begin{aligned} r(t) &= f(t,t) \\ &= r(0) + \int_0^t \left(\sigma(s,t)\int_s^t \sigma(s,u)\,\mathrm{d}u\right)\mathrm{d}s + \int_0^t \sigma(s,t)\,\mathrm{d}W^{\mathbb{Q}}(s) \end{aligned}$$

Market models

1. Libor market model (LMM)
   - Under the forward probability measure $\mathbb{Q}^{\star}(T_{i+1})$, the Libor rate $L_i(t) = L(t, T_i, T_{i+1})$ is a martingale:

$$\mathrm{d}L_i(t) = \gamma_i(t) L_i(t)\, \mathrm{d}W_i^{\mathbb{Q}^{\star}(T_{i+1})}(t)$$

   - We can use the Black formula to price caplets and floorlets where the volatility σ_i is defined by:

$$\sigma_i^2 = \frac{1}{T_i - t}\int_t^{T_i} \gamma_i^2(s)\, \mathrm{d}s$$

2. Swap market model (SMM)
   - We have:

$$\mathrm{d}\,\mathrm{Sw}(t) = \eta(t)\,\mathrm{Sw}(t)\, \mathrm{d}W^{\mathbb{Q}^{\star}}(t)$$

The uncertain volatility model (UVM)

We recall that:

$$V(T) - f(S(T)) = \frac{1}{2}\int_0^T e^{r(T-t)} \boldsymbol{\Gamma}(t)\left(\Sigma^2(T,K) - \sigma^2(t)\right) S^2(t)\,\mathrm{d}t$$

If we assume that $\sigma(t) \in [\sigma^-, \sigma^+]$, we obtain a simple rule for achieving a positive P&L:

- if $\boldsymbol{\Gamma}(t) \geq 0$, we have to hedge the portfolio by considering an implied volatility that is equal to the upper bound σ^+;
- if $\boldsymbol{\Gamma}(t) \leq 0$, we set the implied volatility to the lower bound σ^-.

$\Rightarrow$ This rule is valid if the gamma of the option is always positive or negative, that is when the payoff is convex

How to extend this rule when the gamma can change its sign during the life of the option?

The uncertain volatility model (UVM)

- We assume that:

$$\mathrm{d}S(t) = r(t) S(t) \,\mathrm{d}t + \sigma(t) S(t) \,\mathrm{d}W^{\mathbb{Q}}(t)$$

where:

$$\sigma^{-} \leq \sigma(t) \leq \sigma^{+}$$

- Let $V(t, S(t))$ be the option price, whose payoff is $f(S(T))$. $V(t, S(t))$ is bounded:

$$V^{-}(t, S(t)) \leq V(t, S(t)) \leq V^{+}(t, S(t))$$

where $V^{-}(t, S(t)) = \inf_{\mathbb{Q}(\sigma)} \mathbb{E}^{\mathbb{Q}(\sigma)}\left[\exp\left(-\int_t^T r(s)\,\mathrm{d}s\right) f(S(T))\right]$, $V^{+}(t, S(t)) = \sup_{\mathbb{Q}(\sigma)} \mathbb{E}^{\mathbb{Q}(\sigma)}\left[\exp\left(-\int_t^T r(s)\,\mathrm{d}s\right) f(S(T))\right]$ and $\mathbb{Q}(\sigma)$ denotes all the probability measures

- We can then show that V^{-} and V^{+} satisfy the HJB equation:

$$\sup_{\sigma^{-}\leq\sigma(t)\leq\sigma^{+}} / \inf \left(\frac{1}{2}\sigma^2(t) S^2 \frac{\partial^2 V(t,S)}{\partial S^2} + b(t) S \frac{\partial V(t,S)}{\partial S}\right) + \frac{\partial V(t,S)}{\partial t} - r(t) V(t,$$

The uncertain volatility model (UVM)

- Solving the HJB equation is equivalent to solve the modified Black-Scholes PDE:

$$\left\{ \begin{array}{l} \frac{1}{2}\sigma^2 \left(\boldsymbol{\Gamma}\left(t,S\right)\right) S^2 \partial_S^2 V\left(t,S\right) + b\left(t\right) S \partial_S V\left(t,S\right) + \partial_t V\left(t,S\right) - r\left(t\right) V\left(t,S\right) = \\ V\left(T, S\left(T\right)\right) = f\left(S\left(T\right)\right) \end{array} \right.$$

where:

$$\sigma\left(x\right) = \left\{ \begin{array}{ll} \sigma^+ & \text{if } x \geq 0 \\ \sigma^- & \text{if } x < 0 \end{array} \right. \qquad \text{for } V\left(t, S\left(t\right)\right) = V^+\left(t, S\left(t\right)\right)$$

and:

$$\sigma\left(x\right) = \left\{ \begin{array}{ll} \sigma^- & \text{if } x > 0 \\ \sigma^+ & \text{if } x \leq 0 \end{array} \right. \qquad \text{for } V\left(t, S\left(t\right)\right) = V^-\left(t, S\left(t\right)\right)$$

The uncertain volatility model (UVM)

- Let u_i^m be the numerical solution of $V(t_m, S_i)$. At each iteration m, we approximate the gamma coefficient by the central difference method:

$$\mathbf{\Gamma}(t_m, S_i) \simeq \frac{u_{i+1}^m - 2u_i^m + u_{i+1}^m}{h^2}$$

By assuming that:

$$\operatorname{sign}(\mathbf{\Gamma}(t_m, S_i)) \approx \operatorname{sign}(\mathbf{\Gamma}(t_{m+1}, S_i))$$

we can compute the values taken by $\sigma(\mathbf{\Gamma}(t, S))$ and solve the PDE for the next iteration $m+1$

The uncertain volatility model (UVM)

- If we consider the European call option, we have $\mathbf{\Gamma}(t, S) > 0$, meaning that:

$$V^{+}(t, S(t)) = C_{\mathrm{BS}}\left(t, S(t), \sigma^{+}\right)$$

and:

$$V^{-}(t, S(t)) = C_{\mathrm{BS}}\left(t, S(t), \sigma^{-}\right)$$

where $C_{\mathrm{BS}}(t, S, \sigma)$ is the Black-Scholes price at time t when the underlying price is equal to S and the implied volatility is equal to Σ. Then, the worst-case scenario occurs when the volatility $\sigma(t)$ reaches the upper bound σ^{+}

The uncertain volatility model (UVM)

Example #6

We consider a double KOC barrier option:

$$f_{\mathrm{Barrier}}\left(S\left(T\right)\right) = \mathbb{1}\left\{S\left(t\right) \in \left[L, H\right], t \in \mathcal{T}\right\} \cdot f_{\mathrm{Vanilla}}\left(S\left(T\right)\right)$$

with the following parameters: $K = 100$, $L = 80$, $H = 120$, $T = 1$, $b = 5\%$ and $r = 5\%$. We assume that the volatility $\sigma\left(t\right)$ lies in the range of 15% and 25%. We assume a continuous barrier $\mathcal{T} = [0, 1]$ and a window barrier $\mathcal{T} = [0.25, 0.75]$.

The uncertain volatility model (UVM)

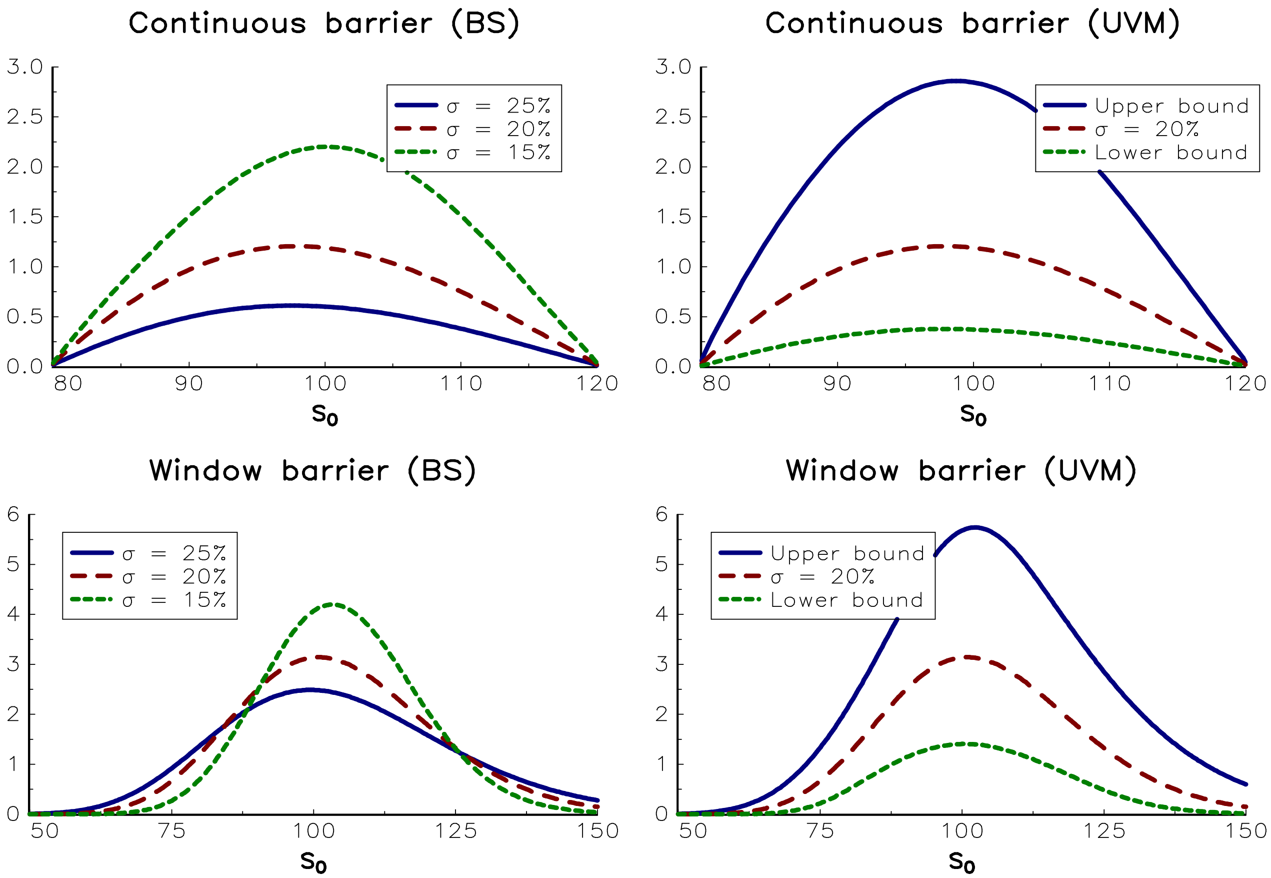


Figure: Comparing BS and UVM prices of the double KOC barrier option

The shifted log-normal model

This model assumes that the asset price $S(t)$ is a linear transformation of a log-normal random variable $X(t)$:

$$S(t) = \alpha(t) + \beta(t) X(t)$$

where $\beta(t) \geq 0$. Then, the payoff of the European call option is:

$$\begin{aligned} f(S(T)) &= (S(T) - K)^+ \\ &= (\alpha(T) + \beta(T) X(T) - K)^+ \\ &= \beta(T) \left(X(T) - \frac{K - \alpha(T)}{\beta(T)} \right)^+ \end{aligned}$$

This type of approach is interesting because the pricing of options can then be done using the Black-Scholes formula:

$$\mathcal{C}(0, S_0) = \beta(T) C_{\mathrm{BS}} \left(X_0, \frac{K - \alpha(T)}{\beta(T)}, \sigma_X, T, b_X, r \right)$$

where b_X and σ_X are the drift and diffusion coefficients of $X(t)$ under the risk-neutral probability measure $\mathbb{Q}$

The fixed-strike parametrization

Let us suppose that:

$$S(t) = \alpha + \beta \exp\left(\left(b^{\mathbb{Q}}(t) - \frac{1}{2}\sigma^2\right)t + \sigma W^{\mathbb{Q}}(t)\right)$$

We have $S_0 = \alpha + \beta$ meaning that:

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

Let b the cost-of-carry parameter of the asset. Under the risk-neutral probability measure, the martingale condition is:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-bt}S(t) \mid \mathcal{F}_0\right] = S_0$$

The fixed-strike parametrization

Since we have $\mathbb{E}^{\mathbb{Q}}\left[S\left(t\right)\right]=\alpha+\left(S_0-\alpha\right)e^{b^{\mathbb{Q}}(t)t}$, we deduce that the no-arbitrage condition implies that:

$$\alpha+\left(S_0-\alpha\right)e^{b^{\mathbb{Q}}(t)t}=S_0e^{bt}\Leftrightarrow b^{\mathbb{Q}}\left(t\right)=\frac{1}{t}\ln\left(\frac{S_0e^{bt}-\alpha}{S_0-\alpha}\right)$$

The payoff of the European call option is:

$$f\left(S\left(T\right)\right)=\left(S\left(T\right)-K\right)^+=\left(\left(S\left(T\right)-\alpha\right)-\left(K-\alpha\right)\right)^+$$

We deduce that the price of the option is given by:

$$\mathcal{C}\left(0,S_0\right)=C_{\mathrm{BS}}\left(S_0-\alpha,K-\alpha,\sigma,T,b^{\mathbb{Q}}\left(T\right),r\right)$$

The fixed-strike parametrization

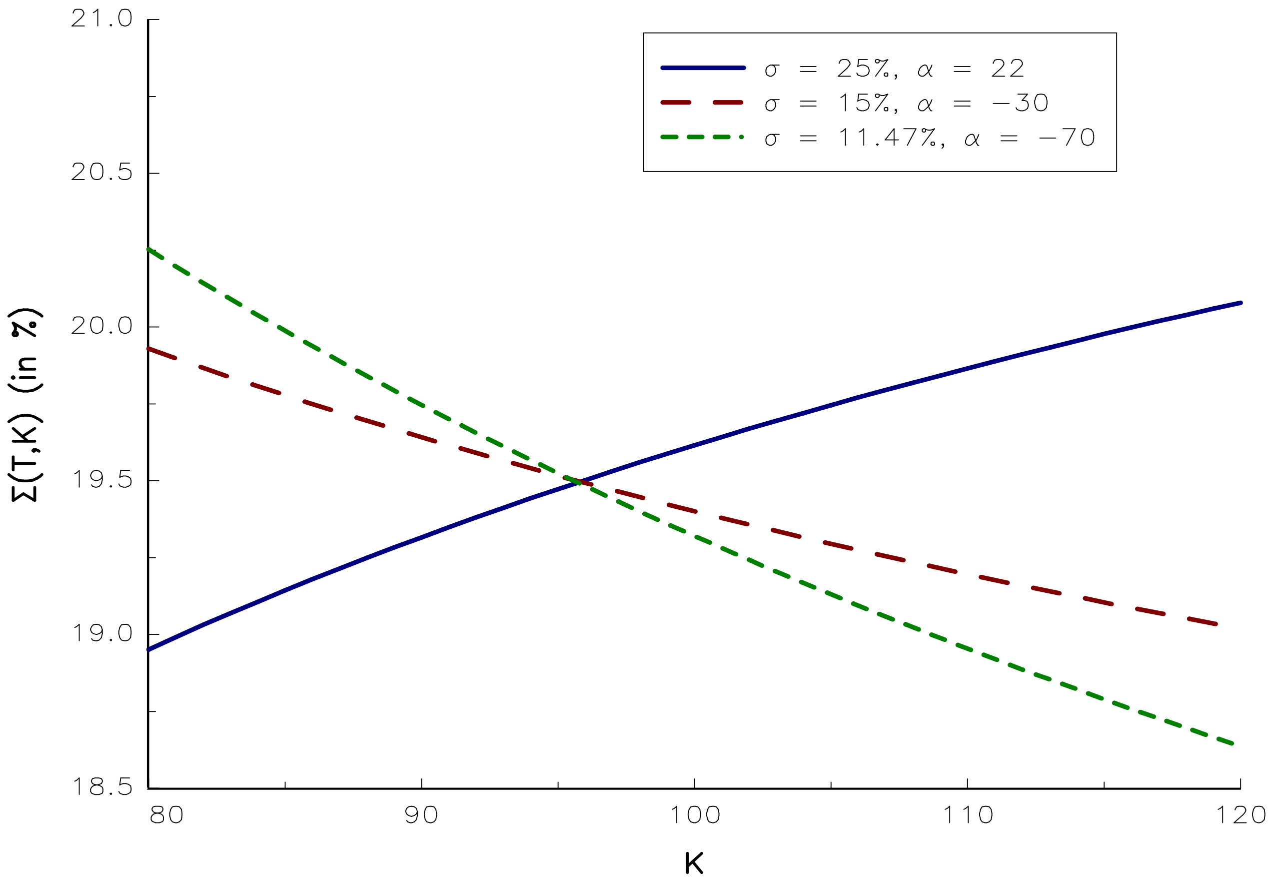


Figure: Volatility skew generated by the SLN model (fixed-strike parametrization)

The floating-strike parametrization

Let us now suppose that:

$$S(t) = \alpha e^{\varphi t} + \beta e^{\left(b - \frac{1}{2}\sigma^2\right)t + \sigma W^{\mathbb{Q}}(t)}$$

We have $S_0 = \alpha + \beta$ and $\mathbb{E}^{\mathbb{Q}}[S(t)] = \alpha e^{\varphi t} + \beta e^{bt}$. We deduce that the stochastic process $e^{-bt} S(t)$ is a $\mathcal{F}_t$-martingale if it is equal to:

$$S(t) = \alpha e^{bt} + (S_0 - \alpha) e^{\left(b - \frac{1}{2}\sigma^2\right)t + \sigma W^{\mathbb{Q}}(t)}$$

The payoff of the European call option becomes:

$$f(S(T)) = (S(T) - K)^+ = \left(\left(S(T) - \alpha e^{bT}\right) - \left(K - \alpha e^{bT}\right)\right)^+$$

It follows that the option price is equal to:

$$\mathcal{C}(0, S_0) = C_{\mathrm{BS}}\left(S_0 - \alpha, K - \alpha e^{bT}, \sigma, T, b, r\right)$$

The floating-strike parametrization

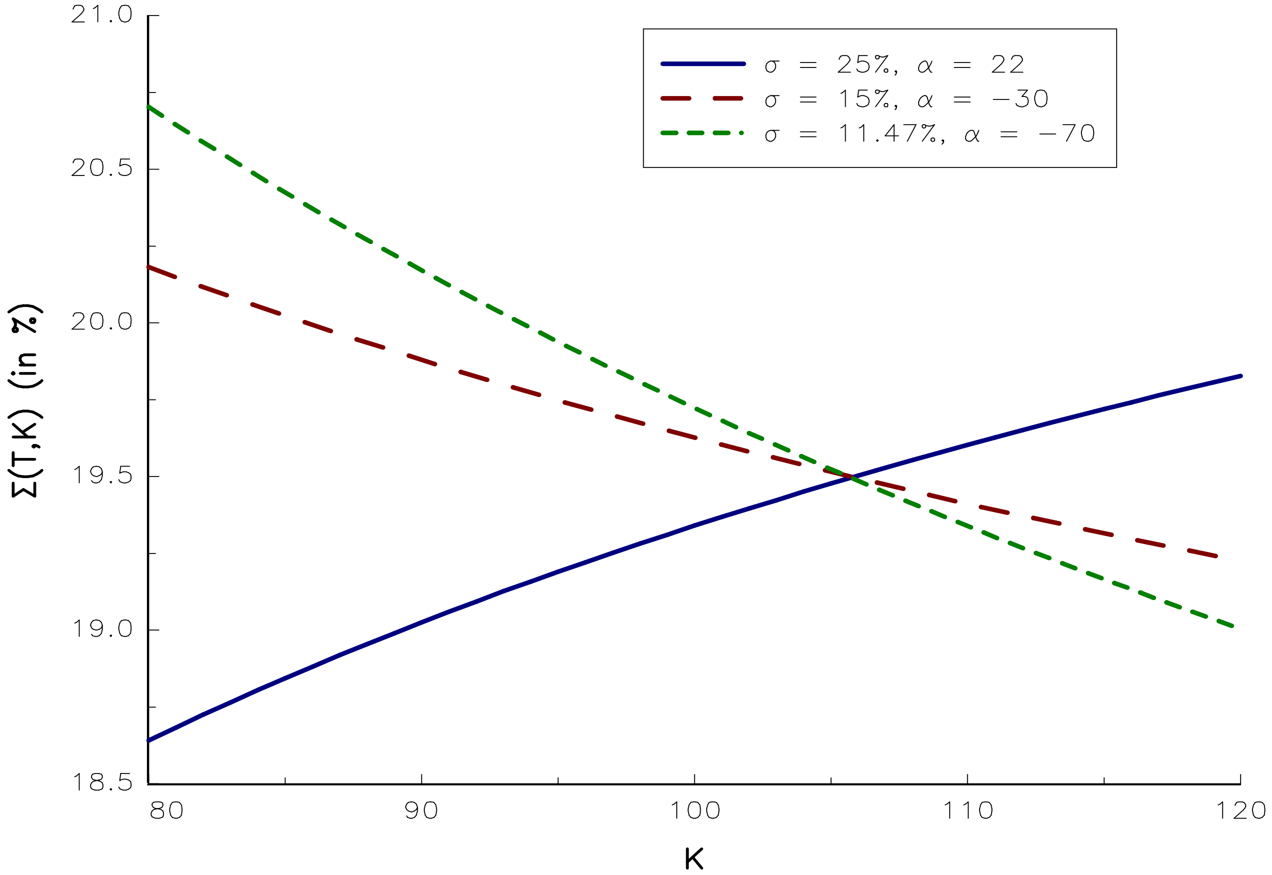


Figure: Volatility skew generated by the SLN model (floating-strike parametrization)

The forward parametrization

If we consider the forward price $F(t)$ instead of the spot price $S(t)$, the two models coincide because we have $b = 0$:

$$\mathrm{d}F(t) = \sigma (F(t) - \alpha)\, \mathrm{d}W^{\mathbb{Q}}(t)$$

and the price of the option is given by the Black formula:

$$\mathcal{C}(0, S_0) = C_{\mathrm{Black}}(F_0 - \alpha, K - \alpha, \sigma, T, r)$$

The forward parametrization

We can prove the following results:

- monotonicity in strike:

$$\operatorname{sign}\left(\frac{\partial\,\Sigma\left(T,K\right)}{\partial\,K}\right)=\operatorname{sign}\alpha$$

- upper and lower bounds:

$$\begin{cases}\Sigma\left(T,K\right)<\sigma & \text{if } \alpha>0\\ \Sigma\left(T,K\right)>\sigma & \text{if } \alpha<0\end{cases}$$

- sharpness of bound:

$$\lim_{K\rightarrow\infty}\Sigma\left(T,K\right)=\sigma$$

- short-expiry behavior:

$$\lim_{T\rightarrow 0}\Sigma\left(T,K\right)=\begin{cases}\dfrac{\sigma\ln\left(F_0/K\right)}{\ln\left(\left(F_0-\alpha\right)/\left(K-\alpha\right)\right)} & \text{if } K\neq F_0\\ \sigma\left(1-\alpha F_0^{-1}\right) & \text{if } K=F_0\end{cases}$$

The forward parametrization

Table: Error of the SLN implied volatility formula (in bps)

K	$(\alpha = 22, \sigma = 25\%)$			$(\alpha = -70, \sigma = 12\%)$		
	1M	1Y	5Y	1M	1Y	5Y
80	1.0	11.1	57.0	−0.9	−12.9	−66.0
90	0.7	10.6	54.1	−1.0	−11.9	−61.4
100	0.9	10.2	51.6	−1.1	−11.3	−57.3
110	1.0	9.7	49.6	−0.8	−10.8	−53.8
120	0.7	9.3	47.7	−0.6	−10.3	−51.3

Mixture of SLN distributions

- One limitation of the SLN model is that it only produces a volatility skew, and not a volatility smile
- The (risk-neutral) probability density function $f(x)$ of the asset price density is given by the mixture of known basic densities:

$$f(x) = \sum_{j=1}^{m} p_j f_j(x)$$

where f_j is the j^{th} basic density, $p_j > 0$ and $\sum_{j=1}^{m} p_j = 1$

- Let $G(S(T))$ be the payoff of an European option. We have:

$$\mathcal{C}(0, S_0) = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT} G(S(T)) \middle| \mathcal{F}_0 \right] = \ldots = \sum_{j=1}^{m} p_j \mathbb{E}^{\mathbb{Q}_j}\left[e^{-rT} G(S(T)) \middle| \mathcal{F}_0 \right]$$

Mixture of SLN distributions

- If we consider a mixture of two shifted log-normal models, the price of the European call option is equal to:

$$\begin{aligned}\mathcal{C}(0, S_0) &= p \cdot C_{\mathrm{SLN}}(S_0, K, \sigma_1, T, b, r, \alpha_1) + \\ &\quad (1-p) \cdot C_{\mathrm{SLN}}(S_0, K, \sigma_2, T, b, r, \alpha_2)\end{aligned}$$

  where C_{SLN} is the formula of the SLN model
- The model has five parameters: σ_1, σ_2, α_1, α_2 and p

Mixture of SLN distributions

Example #7

We consider a calibration set of five options, whose strike and implied volatilities are equal to:

K_j	80	90	100	110	120
$\Sigma(1, K_j)$	21%	19%	18.25%	18.5%	19%

The current value of the asset price is equal to 100, the maturity of options is one year, the cost-of-carry parameter is set to 0 and the interest rate is 5%

The parameters are estimated by minimizing the weighted least squares:

$$\min \sum_{j=1}^{n} w_j \left(\hat{C}_j - C_{\mathrm{SLN}}\left(S_0, K_j, \sigma_1, \sigma_2, T_j, b, r, \alpha_1, \alpha_2, p\right) \right)^2$$

where $\hat{C}_j = C_{\mathrm{BS}}\left(S_0, K_j, \Sigma\left(T_j, K_j\right), T_j, b, r\right)$ and w_j is the weight of the j^{th} option

Mixture of SLN distributions

We consider three parameterizations:

- (#1) the weights w_j are uniform, and we impose that $\alpha_1 = \alpha_2$ and $p = 50\%$
- (#2) the weights w_j are uniform, and p is set to 25%
- (#3) the weights w_j are inversely proportional to option prices $\hat{C}_j$, and p is set to 50%

Table: Calibrated parameters of the mixed SLN model

Model	#1	#2	#3
σ_1	16.5%	8.2%	10.2%
σ_2	7.3%	17.2%	21.7%
α_1	-53.3	-289.7	-145.2
α_1	-53.3	19.6	47.4
p	50.0%	25.0%	50.0%

Mixture of SLN distributions

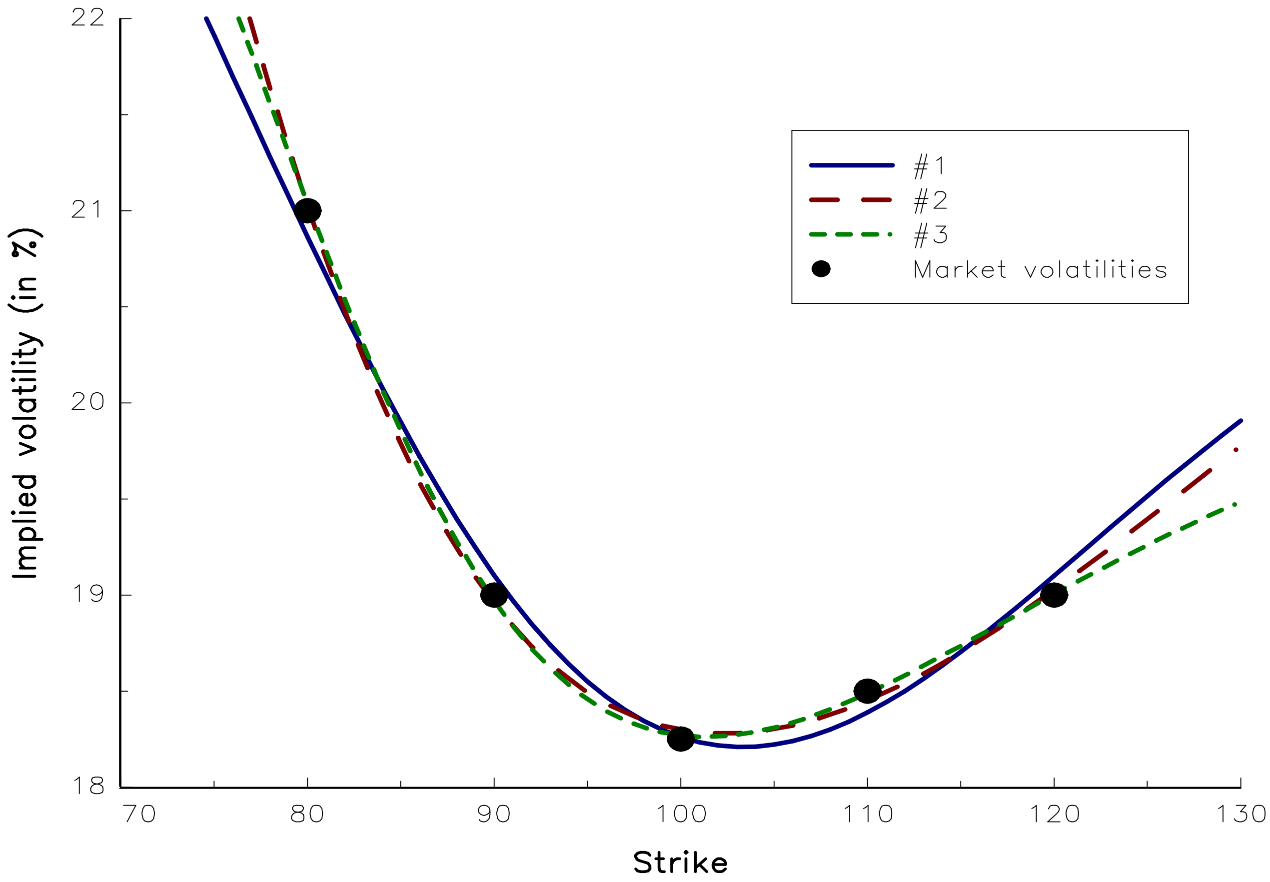


Figure: Implied volatility (in %) of calibrated mixed SLN models

Local volatility model

- We assume that:

$$\mathrm{d}S(t) = bS(t)\,\mathrm{d}t + \sigma(t, S(t))\,S(t)\,\mathrm{d}W^{\mathbb{Q}}(t)$$

- $\sigma(t, S) \Rightarrow \Sigma(T, K)$
- $\Sigma(T, K) \Rightarrow \sigma(t, S)$ (Dupire model)
- Relationship with the Breeden-Litzenberger model

The Fokker-Planck equation

The risk-neutral probability density function $q_t(T, S)$ of the asset price $S(T)$ satisfies the forward Chapman-Kolmogorov equation:

$$\frac{\partial\, q_t(T, S)}{\partial\, T} = -\frac{\partial\, [bSq_t(T, S)]}{\partial\, S} + \frac{1}{2}\frac{\partial^2\, \left[\sigma^2(T, S)\, S^2 q_t(T, S)\right]}{\partial\, S^2}$$

The initial condition is:

$$q_t(t, S) = \mathbb{1}\{S = S_t\}$$

where S_t is the value of $S(t)$ that is known at time t

The Breeden-Litzenberger formulas

We have:

$$\begin{aligned}
\mathcal{C}_t(T,K) &= e^{-r(T-t)}\int_K^\infty (S-K)\, q_t(T,S)\,\mathrm{d}S \\
\frac{\partial\,\mathcal{C}_t(T,K)}{\partial\,K} &= -e^{-r(T-t)}\int_K^\infty q_t(T,S)\,\mathrm{d}S \\
\frac{\partial^2\,\mathcal{C}_t(T,K)}{\partial\,K^2} &= e^{-r(T-t)} q_t(T,K) \\
\frac{\partial\,\mathcal{C}_t(T,K)}{\partial\,T} &= -r\mathcal{C}_t(T,K) + e^{-r(T-t)}\int_K^\infty (S-K)\,\frac{\partial\,q_t(T,S)}{\partial\,T}\,\mathrm{d}S
\end{aligned}$$

Derivation of the forward equation

We deduce that:

$$\frac{\partial \mathcal{C}_t(T,K)}{\partial T} = -r\mathcal{C}_t(T,K) + e^{-r(T-t)}\mathcal{I}$$

where:

$$\mathcal{I} = \frac{1}{2}\sigma^2(T,K)K^2 q_t(T,K) + be^{r(T-t)}\left(\mathcal{C}_t(T,K) - K\frac{\partial \mathcal{C}_t(T,K)}{\partial K}\right)$$

It follows that:

$$\begin{aligned}\frac{\partial \mathcal{C}_t(T,K)}{\partial T} = & -r\mathcal{C}_t(T,K) + \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 \mathcal{C}_t(T,K)}{\partial K^2} + \\ & b\left(\mathcal{C}_t(T,K) - K\frac{\partial \mathcal{C}_t(T,K)}{\partial K}\right)\end{aligned}$$

Derivation of the forward equation

We conclude that:

$$\frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 \mathcal{C}_t(T,K)}{\partial K^2} - bK\frac{\partial \mathcal{C}_t(T,K)}{\partial K} - \frac{\partial \mathcal{C}_t(T,K)}{\partial T} + (b-r)\mathcal{C}_t(T,K) = 0$$

Differences between backward and forward PDE approaches

The backward PDE is:

$$\left\{ \begin{array}{l} \frac{1}{2}\sigma^2 (t, S)\, S^2 \partial_S^2 V(t, S) + bS\partial_S V(t, S) + \partial_t V(t, S) - rV(t, S) = 0 \\ V(T, S(T)) = f(T, S(T), K) \end{array} \right.$$

The forward PDE is:

$$\left\{ \begin{array}{l} \frac{1}{2}\sigma^2 (T, K)\, K^2 \partial_K^2 V(T, K) - bK\partial_K V(T, K) - \partial_T V(T, K) + (b - r)\, V(T, K) = \\ V(t, K) = f(t, S_t, K) \end{array} \right.$$

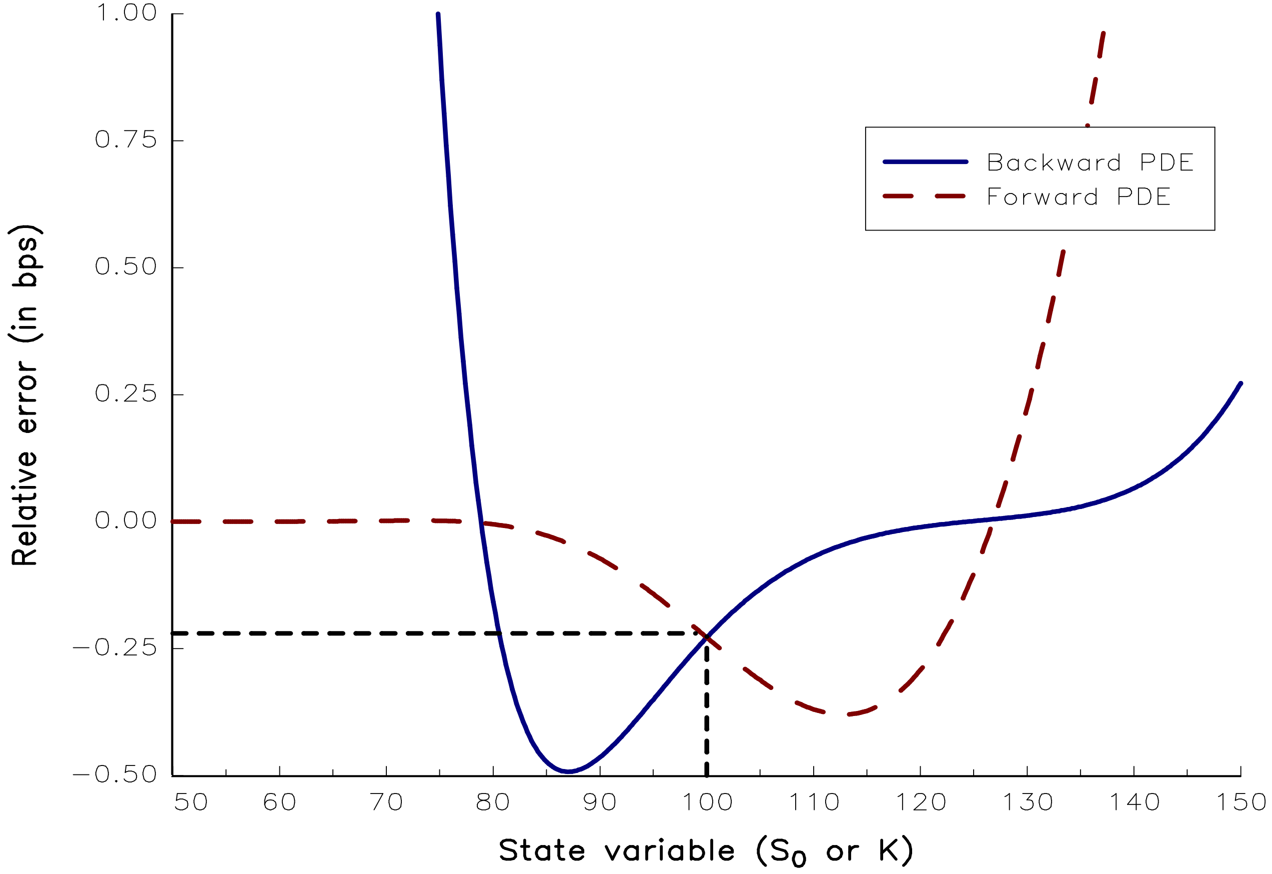
Differences between backward and forward PDE approaches

- In the backward formulation, the state variables are t and S, whereas the fixed variables are T and K
- In the forward formulation, the state variables are T and K, whereas the fixed variables are the current time t and the current asset price S_t
- The backward PDE approach suggests that we can hedge the option using a dynamic portfolio of the underlying asset
- The forward PDE approach suggests that we can hedge the option using a static portfolio of call and put options

Differences between backward and forward PDE approaches

- We consider the pricing of an European call option with the following parameters: $S_0 = 100$, $K = 100$, $\sigma(t, S) = 20\%$, $T = 0.5$, $b = 2\%$ and $r = 5\%$
- In the case of the backward PDE, we consider the usual boundary conditions:
$$\begin{cases} \mathcal{C}(t, S) = 0 \\ \partial_S \mathcal{C}(t, +\infty) = 1 \end{cases}$$
- For the forward PDE, the boundary conditions are:
$$\begin{cases} \partial_K \mathcal{C}(T, 0) = -1 \\ \mathcal{C}(T, +\infty) = 0 \end{cases}$$
- The backward and forward PDEs are solved using the Crank-Nicholson scheme

Differences between backward and forward PDE approaches



Duality between local volatility and implied volatility

We have:

$$\sigma^2(T,K) = 2\frac{bK\partial_K \mathcal{C}(T,K) + \partial_T \mathcal{C}(T,K) - (b-r)\mathcal{C}(T,K)}{K^2 \partial_K^2 \mathcal{C}(T,K)}$$

Duality between local volatility and implied volatility

We can show:

$$\sigma(T,K)=\sqrt{\frac{A(T,K)}{B(T,K)}}$$

where:

$$\begin{aligned}A(T,K) &= \Sigma^2(T,K)+2bKT\Sigma(T,K)\,\partial_K\Sigma(T,K)+\\&\quad 2T\Sigma(T,K)\,\partial_T\Sigma(T,K)\end{aligned}$$

and:

$$\begin{aligned}B(T,K) &= 1+2K\sqrt{T}d_1\partial_K\Sigma(T,K)+K^2T\Sigma(T,K)\,\partial_K^2\Sigma(T,K)+\\&\quad K^2Td_1d_2\left(\partial_K\Sigma(T,K)\right)^2\end{aligned}$$

Duality between local volatility and implied volatility

Example #8

We assume that the implied volatility is equal to:

$$\Sigma\left(T,K\right)=\Sigma_0+\alpha\left(S_0-K\right)^2$$

where $\Sigma_0 = 20\%$, $\alpha = 1$ bp, $S_0 = 100$ and $b = 5\%$.

Duality between local volatility and implied volatility

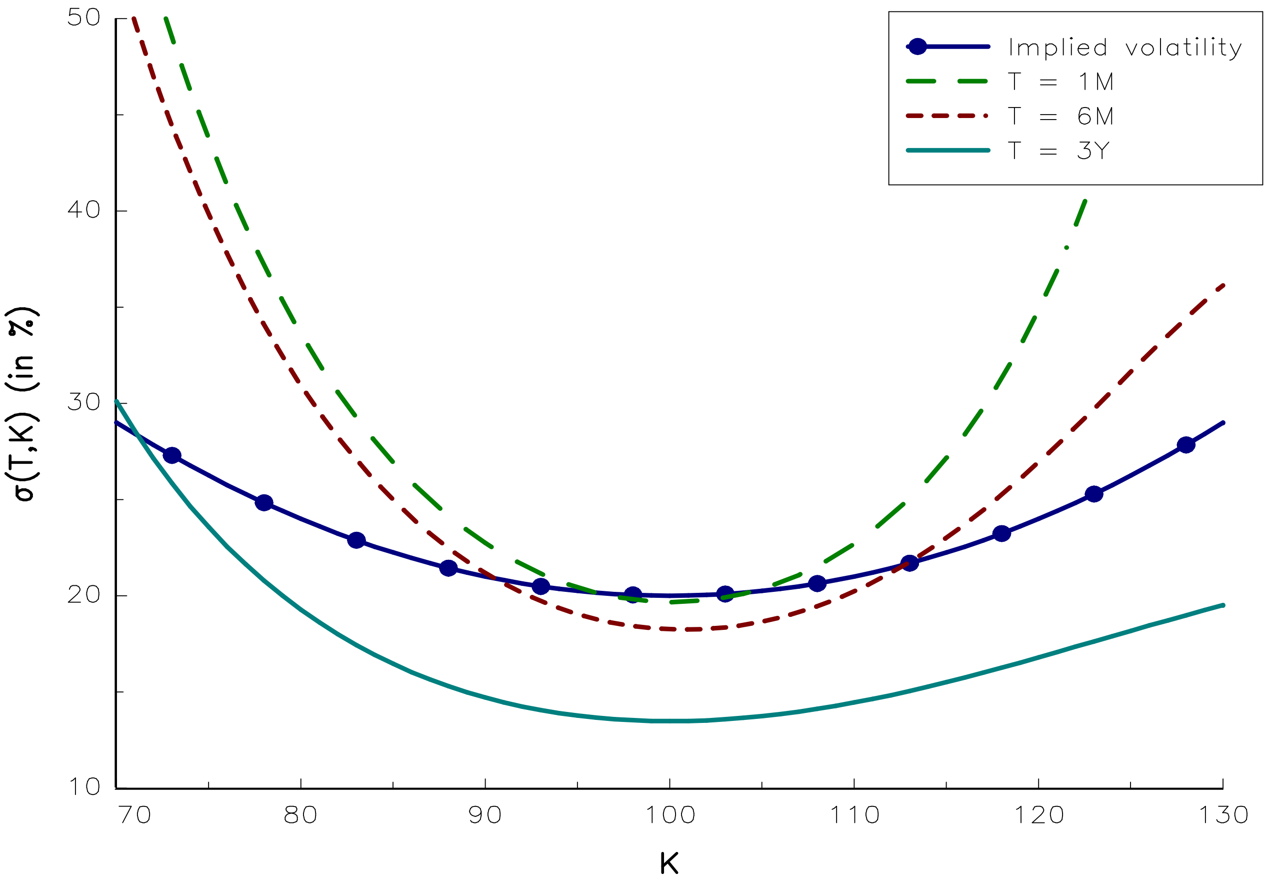


Figure: Calibrated local volatility $\sigma(T, S)$ (in %)

Duality between local volatility and implied volatility

If there is no skew, the local volatility function does not depend on the strike:

$$\sigma^2(T) = \Sigma^2(T) + 2T\Sigma(T)\frac{\partial\,\Sigma(T)}{\partial\,T}$$

We notice that:

$$\sigma^2(T) = \Sigma^2(T) + 2T\Sigma(T)\frac{\partial\,\Sigma(T)}{\partial\,T} = \frac{\partial\,T\Sigma^2(T)}{\partial\,T}$$

or:

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

The implied variance is then the time series average of the local variance

Duality between local volatility and implied volatility

- Let x be the log-moneyness:

$$x = \varphi(T, K) = \ln \frac{S_0}{K} + bT$$

- We introduce the functions $\tilde{\Sigma}$ and $\tilde{\sigma}$ such that $\Sigma(T, K) = \tilde{\Sigma}(T, \varphi(T, K))$ and $\sigma(T, K) = \tilde{\sigma}(T, \varphi(T, K))$
- We can show that the implied volatility is the harmonic mean of the local volatility:

$$\frac{1}{\tilde{\Sigma}(0, x)} = \int_0^1 \frac{\mathrm{d}y}{\tilde{\sigma}(0, xy)}$$

- It follows that:

$$\frac{\partial \tilde{\Sigma}(0, 0)}{\partial x} = \frac{1}{2} \frac{\partial \tilde{\sigma}(0, 0)}{\partial x}$$

- The ATM slope of the implied volatility near expiry is equal to one half the slope of the local volatility

Dupire model in practice

- Time interpolation (e.g., linear interpolation of the total implied variance $\upsilon(T,K) = T\Sigma^2(T,K)$)
- Non-parametric interpolation (e.g., cubic spline interpolation)
- Parametric interpolation (e.g., SVI parametrization)

Dupire model in practice

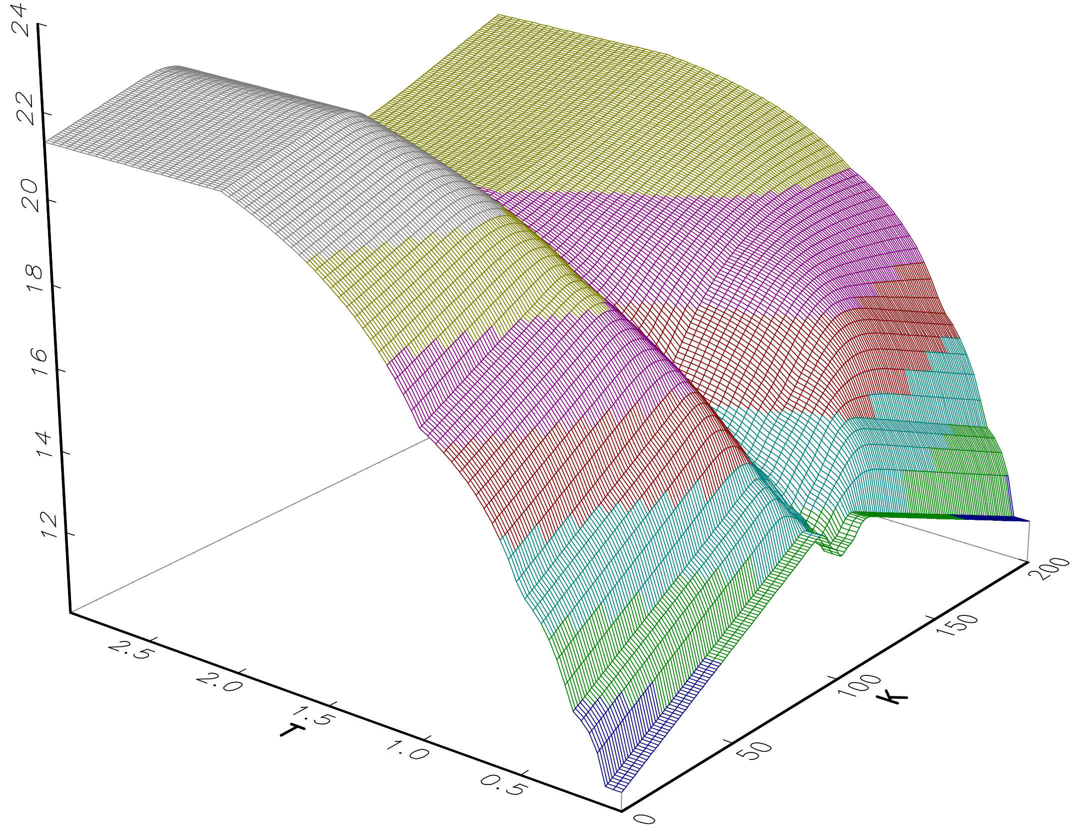


Figure: Implied volatility surface $\Sigma(T, K)$ (in %)

Dupire model in practice

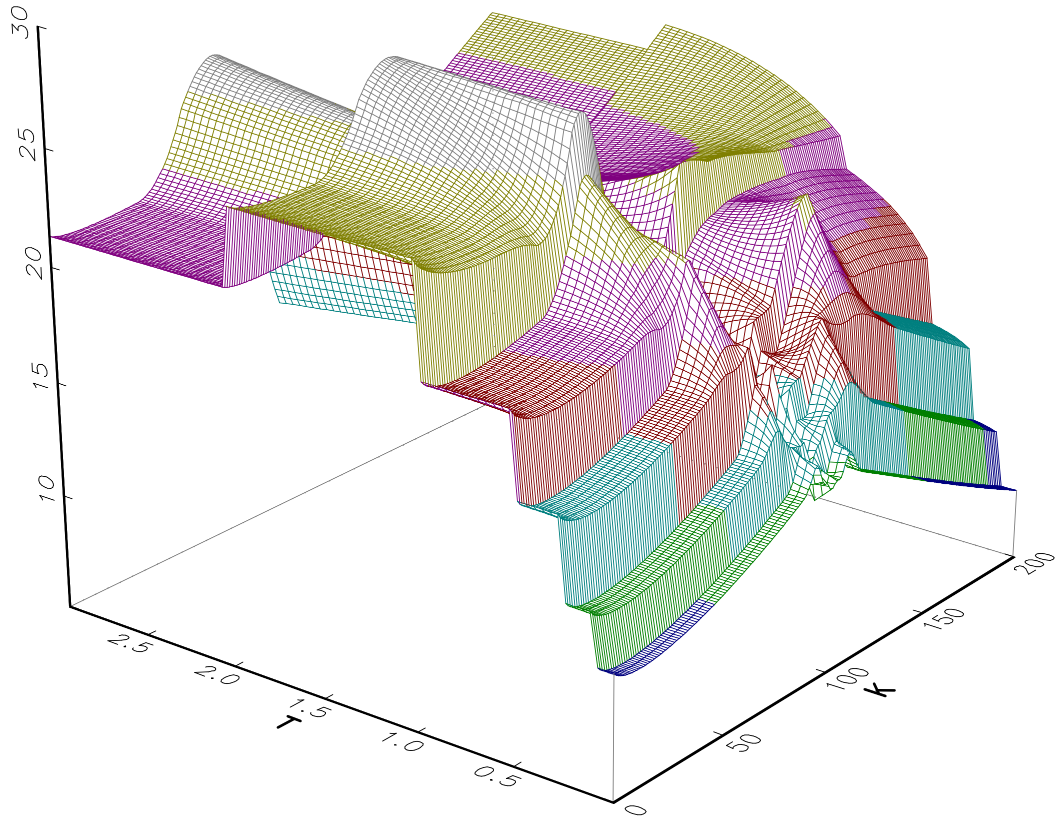


Figure: Local volatility surface $\sigma(T, K)$ (in %)

Hedging coefficients

- The delta of the option is:

$$\boldsymbol{\Delta} \approx \frac{V(T,K,S_t+\varepsilon) - V(T,K,S_t-\varepsilon)}{2\varepsilon}$$

- The gamma of the option is:

$$\boldsymbol{\Gamma} \approx \frac{V(T,K,S_t+\varepsilon) - 2V(T,K,S_t) + V(T,K,S_t-\varepsilon)}{\varepsilon^2}$$

- The vega is the sensitivity of the price to a parallel shift of $\Sigma(T,K,S_t)$:

$$\boldsymbol{\upsilon} = \frac{V'(T,K,S_t) - V(T,K,S_t)}{\varepsilon'}$$

where $V'(T,K,S_t)$ is the option price obtained when the implied volatility surface is $\Sigma(T,K,S_t) + \varepsilon'$

Hedging coefficients

"Market smiles and skews are usually managed by using local volatility models a la Dupire. We discover that the dynamics of the market smile predicted by local vol models is opposite of observed market behavior: when the price of the underlying decreases, local vol models predict that the smile shifts to higher prices; when the price increases, these models predict that the smile shifts to lower prices. Due to this contradiction between model and market, delta and vega hedges derived from the model can be unstable and may perform worse than naive Black-Scholes' hedges" (Hagan et al., 2002, page 84).

Application to exotic options

- We assume that S_0, $b = 5\%$ and r
- The option parameters are $K = 100$, $L = 90$ and $H = 115$
- The maturity is set to one year

Table: Barrier option pricing with the local volatility model

Option	Payoff	LV	BS-PDE Σ_1	BS-PDE Σ_2	BS-RR Σ_1	BS-RR Σ_2
Call	$(S(T) - K)^+$	8.85	8.96	8.78	8.96	8.78
Put	$(K - S(T))^+$	3.97	4.08	3.90	4.08	3.90
DOC	$\mathbb{1}\{S(t) > L\} \cdot (S(T) - K)^+$	7.98	8.14	8.05	8.11	8.02
DOP	$\mathbb{1}\{S(t) > L\} \cdot (K - S(T))^+$	0.26	0.27	0.28	0.25	0.27
UOC	$\mathbb{1}\{S(t) < H\} \cdot (S(T) - K)^+$	0.99	0.88	0.94	0.83	0.89
UOP	$\mathbb{1}\{S(t) < H\} \cdot (K - S(T))^+$	3.81	3.90	3.75	3.89	3.74
KOC	$\mathbb{1}\{S(t) \in [L, H]\} \cdot (S(T) - K)^+$	0.65	0.56	0.64	0.52	0.59
KOP	$\mathbb{1}\{S(t) \in [L, H]\} \cdot (K - S(T))^+$	0.20	0.20	0.22	0.19	0.21
BCC	$\mathbb{1}\{S(T) \geq K\}$	0.58	0.56	0.57	0.56	0.57
BCP	$\mathbb{1}\{S(T) \leq K\}$	0.37	0.39	0.38	0.39	0.38

Stochastic volatility models

We assume that the joint dynamics of the spot price $S(t)$ and the stochastic volatility $\sigma(t)$ is:

$$\begin{cases} \mathrm{d}S(t) = \mu(t) S(t) \,\mathrm{d}t + \sigma(t) S(t) \,\mathrm{d}W_1(t) \\ \mathrm{d}\sigma(t) = \zeta(\sigma(t)) \,\mathrm{d}t + \xi(\sigma(t)) \,\mathrm{d}W_2(t) \end{cases}$$

where $\mathbb{E}[W_1(t) W_2(t)] = \rho t$

Stochastic volatility models

The fundamental pricing equation is:

$$\begin{aligned} \frac{1}{2}\sigma^2 S^2 \partial_S^2 V(t,S,\sigma) + \rho\sigma S\xi(\sigma)\,\partial_{S,\sigma}^2 V(t,S,\sigma) + \frac{1}{2}\xi^2(\sigma)\,\partial_\sigma^2 V(t,S,\sigma) & \\ +(\mu - \lambda_S\sigma) S\partial_S V(t,S,\sigma) + (\zeta(\sigma) - \lambda_\sigma\xi(\sigma))\,\partial_\sigma V(t,S,\sigma) & \\ +\partial_t V(t,S,\sigma) - rV(t,S,\sigma) & = 0 \end{aligned}$$

where $V(t,S,\sigma)$ is the price of the contingent claim, $V(T,S(T)) = f(S(T))$ and $f(S(T))$ is the option payoff

Stochastic volatility models

- The market price of the spot risk $W_1(t)$ is:

$$\lambda_S(t) = \frac{\mu(t) - b(t)}{\sigma(t)}$$

- We introduce the function $\zeta'(y)$:

$$\zeta'(\sigma(t)) = \zeta(\sigma(t)) - \lambda_\sigma(t)\,\xi(\sigma(t))$$

- The PDE becomes:

$$\begin{aligned} \frac{1}{2}\sigma^2 S^2 \partial_S^2 V(t,S,\sigma) + \rho\sigma S\xi(\sigma)\,\partial_{S,\sigma}^2 V(t,S,\sigma) + \frac{1}{2}\xi^2(\sigma)\,\partial_\sigma^2 V(t,S,\sigma) & \\ +bS\partial_S V(t,S,\sigma) + \zeta'(\sigma)\,\partial_\sigma V(t,S,\sigma) + \partial_t V(t,S,\sigma) - rV(t,S,\sigma) &= 0 \end{aligned}$$

Stochastic volatility models

- Using the Girsanov theorem, we deduce that the risk-neutral dynamics is:

$$\left\{ \begin{array}{l} \mathrm{d}S(t) = b(t) S(t) \,\mathrm{d}t + \sigma(t) S(t) \,\mathrm{d}W_1^{\mathbb{Q}}(t) \\ \mathrm{d}\sigma(t) = \zeta'(\sigma(t)) \,\mathrm{d}t + \xi(\sigma(t)) \,\mathrm{d}W_2^{\mathbb{Q}}(t) \end{array} \right.$$

- The martingale solution is then equal to:

$$V_0 = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_0^T r(t)\,\mathrm{d}t} f(S(T)) \middle| \mathcal{F}_0 \right]$$

Hedging portfolio

In the case of the Black-Scholes model, delta and vega sensitivities are equal to:

$$\boldsymbol{\Delta}_{\mathrm{BS}} = \frac{\partial\, V_{\mathrm{BS}}\left(S_0, K, \Sigma, T\right)}{\partial\, S_0}$$

and:

$$\boldsymbol{\upsilon}_{\mathrm{BS}} = \frac{\partial\, V_{\mathrm{BS}}\left(S_0, K, \Sigma, T\right)}{\partial\, \Sigma}$$

Hedging portfolio

In the case of the stochastic volatility model, we have:

$$\boldsymbol{\Delta}_{\mathrm{SV}} = \frac{\partial\, V_{\mathrm{SV}}\left(S_0, K, \sigma_0, T\right)}{\partial\, S_0}$$

If we assume that $V_{\mathrm{SV}}\left(S_0, K, \sigma_0, T\right) = V_{\mathrm{BS}}\left(S_0, K, \Sigma_{\mathrm{SV}}\left(T, S_0\right), T\right)$, we obtain:

$$\begin{aligned}\boldsymbol{\Delta}_{\mathrm{SV}} &= \frac{\partial\, V_{\mathrm{BS}}\left(S_0, K, \Sigma_{\mathrm{SV}}, T\right)}{\partial\, S_0} + \frac{\partial\, V_{\mathrm{BS}}\left(S_0, K, \Sigma_{\mathrm{SV}}, T\right)}{\partial\, \Sigma_{\mathrm{SV}}} \cdot \frac{\partial\, \Sigma_{\mathrm{SV}}\left(T, S_0\right)}{\partial\, S_0} \\ &= \boldsymbol{\Delta}_{\mathrm{BS}} + \boldsymbol{\upsilon}_{\mathrm{BS}} \cdot \frac{\partial\, \Sigma_{\mathrm{SV}}\left(T, S_0\right)}{\partial\, S_0}\end{aligned}$$

$\Rightarrow$ The delta of the SV model depends on the BS vega

Generally, we have $\partial_{S_0} \Sigma_{\mathrm{SV}}\left(T, S_0\right) \geq 0$ implying that $\boldsymbol{\Delta}_{\mathrm{SV}} \geq \boldsymbol{\Delta}_{\mathrm{BS}}$

Hedging portfolio

- The natural hedging portfolio should consist in two long/short exposures since we have two risk factors $S(t)$ and $\sigma(t)$
- We can define the vega sensitivity as follows:

$$\boldsymbol{\upsilon}_{\mathrm{SV}} = \frac{\partial\, V_{\mathrm{SV}}\left(S_0, K, \sigma_0, T\right)}{\partial\, \sigma_0}$$

  However, this definition has no interest since the stochastic volatility $\sigma(t)$ cannot be directly or even indirectly trade
- This is why most of traders prefer to use a BS vega:

$$\boldsymbol{\upsilon}_{\mathrm{SV}} = \frac{\partial\, V_{\mathrm{BS}}\left(S_0, K, \Sigma_{\mathrm{SV}}\left(T, S_0\right), T\right)}{\partial\, \Sigma_{\mathrm{SV}}}$$

  The vega is calculated with respect to the implied volatility $\Sigma_{\mathrm{SV}}(T, S_0)$ deduced from the stochastic volatility model

Heston model

We have:

$$\left\{ \begin{array}{l} \mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sqrt{v(t)}S(t)\,\mathrm{d}W_1(t) \\ \mathrm{d}v(t) = \kappa(\theta - v(t))\,\mathrm{d}t + \xi\sqrt{v(t)}\,\mathrm{d}W_2(t) \end{array} \right.$$

where $S(0) = S_0$, $v(0) = v_0$ and $W(t) = (W_1(t), W_2(t))$ is a two-dimensional Wiener process with $\mathbb{E}[W_1(t)\,W_2(t)] = \rho\,t$

Heston model

- The stochastic variance $v(t)$ follows a CIR process: θ is the long-run variance, κ is the mean-reverting parameter and ξ is the volatility of the variance (also called the vovol parameter)
- We have $\sigma(t) = \sqrt{v(t)}$ and:

$$\mathrm{d}\sigma(t) = \left(\left(\frac{\kappa\theta}{2} - \frac{\xi^2}{8}\right)\frac{1}{\sigma(t)} - \frac{1}{2}\kappa\sigma(t)\right)\mathrm{d}t + \frac{1}{2}\xi\,\mathrm{d}W_2(t)$$

The stochastic volatility is then an Ornstein-Uhlenbeck process if we impose $\theta = \xi^2/(4\kappa)$

Heston model

The PDE is:

$$\begin{aligned}\frac{1}{2}vS^2\partial_S^2 V + \rho\xi vS\partial_{S,v}^2 V + \frac{1}{2}\xi^2 v\partial_v^2 V & \\ +bS\partial_S V + (\kappa(\theta - v(t)) - \lambda v)\partial_v V + \partial_t V - rV &= 0\end{aligned}$$

The risk-neutral dynamics is:

$$\left\{\begin{array}{l}\mathrm{d}S(t) = bS(t)\,\mathrm{d}t + \sqrt{v(t)}S(t)\,\mathrm{d}W_1^{\mathbb{Q}}(t) \\ \mathrm{d}v(t) = (\kappa(\theta - v(t)) - \lambda v(t))\,\mathrm{d}t + \xi\sqrt{v(t)}\,\mathrm{d}W_2^{\mathbb{Q}}(t)\end{array}\right.$$

Heston model

The closed-form solutions of European call and put options are:

$$\begin{aligned}\mathcal{C}_0 &= S_0 e^{(b-r)T} P_1 - K e^{-rT} P_2 \\ \mathcal{P}_0 &= S_0 e^{(b-r)T} (P_1 - 1) - K e^{-rT} (P_2 - 1)\end{aligned}$$

where the probabilities P_1 and P_2 satisfy:

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left(\frac{e^{-i\phi \ln K} \varphi_j (S_0, v_0, T, \phi)}{i\phi}\right) \mathrm{d}\phi$$

$$\varphi_j (S_0, v_0, T, \phi) = \exp\left(C_j (T, \phi) + D_j (T, \phi) v_0 + i\phi \ln S_0\right)$$

$$C_j (T, \phi) = ib\phi T + \frac{a_j}{\xi^2}\left((b_j - i\rho\xi\phi + d_j) T - 2 \ln\left(\frac{1 - g_j e^{d_j T}}{1 - g_j}\right)\right)$$

$$D_j (T, \phi) = \frac{b_j - i\rho\xi\phi + d_j}{\xi^2}\left(\frac{1 - e^{d_j T}}{1 - g_j e^{d_j T}}\right)$$

$$g_j = \frac{b_j - i\rho\xi\phi + d_j}{b_j - i\rho\xi\phi - d_j}$$

$$d_j = \sqrt{(i\rho\xi\phi - b_j)^2 - \xi^2 (2iu_j\phi - \phi^2)}$$

where $a_1 = a_2 = \kappa\theta$, $b_1 = \kappa + \lambda - \rho\xi$, $b_2 = \kappa + \lambda$, $u_1 = 1/2$ and $u_2 = -1/2$

Heston model

Example # 9

The parameters are equal to $S_0 = 100$, $b = r = 5\%$, $v_0 = \theta = 4\%$, $\kappa = 0.5$, $\xi = 0.9$ and $\lambda = 0$. We consider the pricing of the European call option, whose maturity is three months.

Heston model

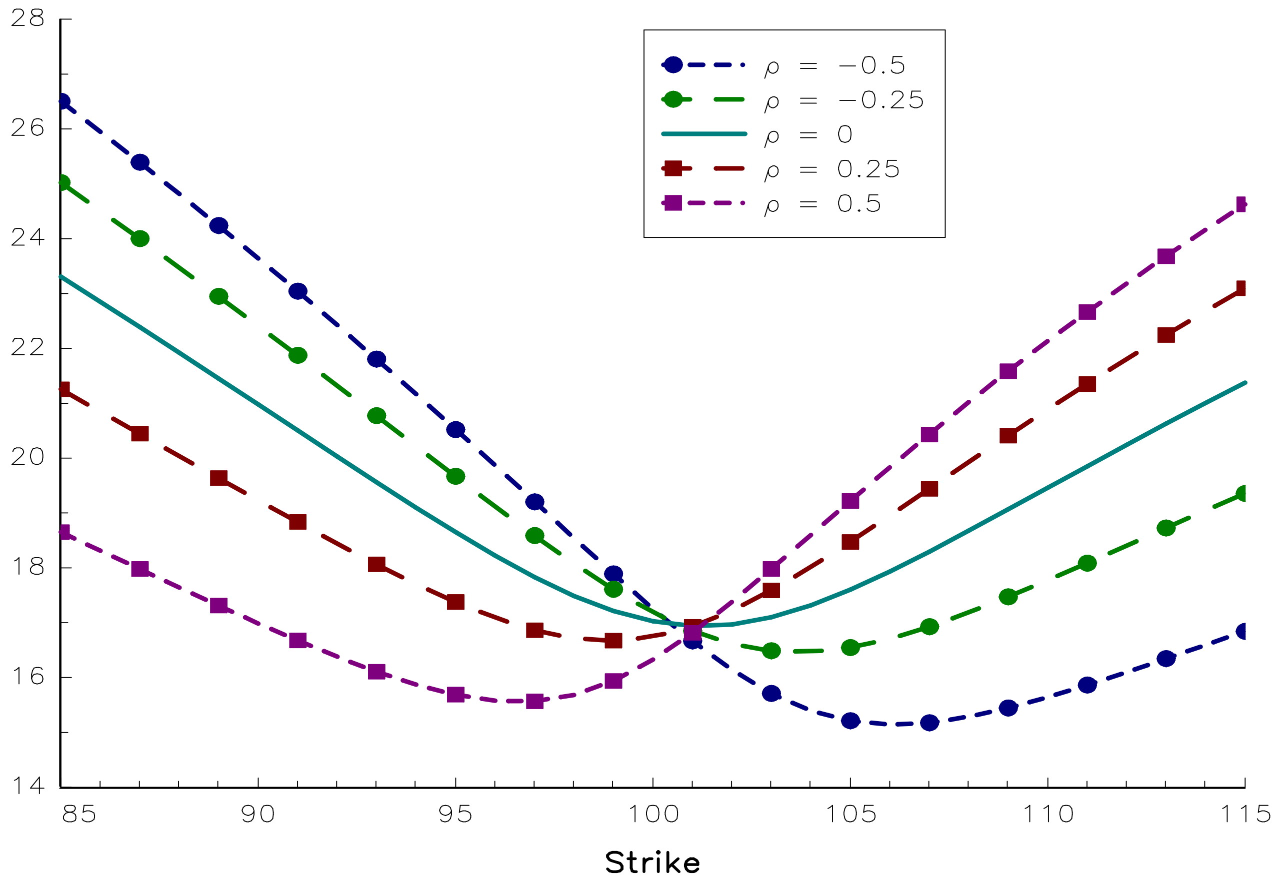


Figure: Implied volatility of the Heston model (in %)

Heston model

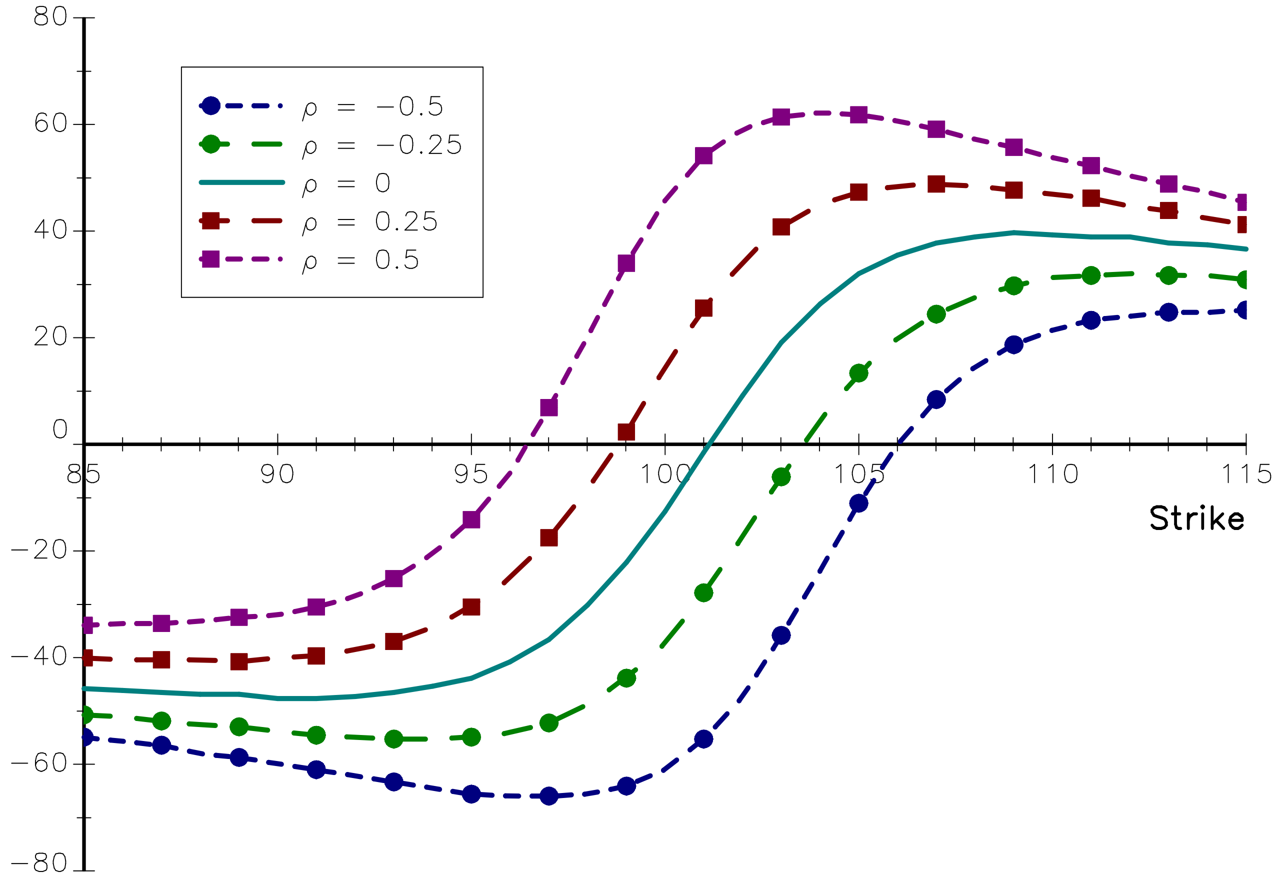


Figure: Skew $\omega(T,K) = \dfrac{\partial \Sigma(T,K)}{\partial K}$ of the Heston model (in bps)

SABR model

The dynamics of the forward rate $F(t)$ is given by:

$$\left\{ \begin{array}{l} \mathrm{d}F(t) = \alpha(t) F(t)^{\beta} \, \mathrm{d}W_1^{\mathbb{Q}}(t) \\ \mathrm{d}\alpha(t) = \nu \alpha(t) \, \mathrm{d}W_2^{\mathbb{Q}}(t) \end{array} \right.$$

where $\mathbb{E}\left[W_1^{\mathbb{Q}}(t) W_2^{\mathbb{Q}}(t)\right] = \rho t$

The model has 4 parameters:

1. α the current value of $\alpha(t)$
2. β the exponent of the forward rate
3. ν the log-normal volatility of $\alpha(t)$
4. ρ the correlation between the two Brownian motions

SABR model

The implied Black volatility is:

$$\Sigma_B(T,K) = \frac{\alpha}{(F_0K)^{(1-\beta)/2}\left(1+\frac{(1-\beta)^2}{24}\ln^2\frac{F_0}{K}+\frac{(1-\beta)^4}{1920}\ln^4\frac{F_0}{K}\right)}\left(\frac{z}{\chi(z)}\right)\cdot$$
$$\left(1+\left(\frac{(1-\beta)^2\alpha^2}{24(F_0K)^{1-\beta}}+\frac{\rho\alpha\nu\beta}{4(F_0K)^{(1-\beta)/2}}+\frac{2-3\rho^2}{24}\nu^2\right)T\right)$$

where:

$$z=\nu\alpha^{-1}(F_0K)^{(1-\beta)/2}\ln\frac{F_0}{K}$$

and:

$$\chi(z)=\ln\left(\sqrt{1-2\rho z+z^2}+z-\rho\right)-\ln(1-\rho)$$

SABR model

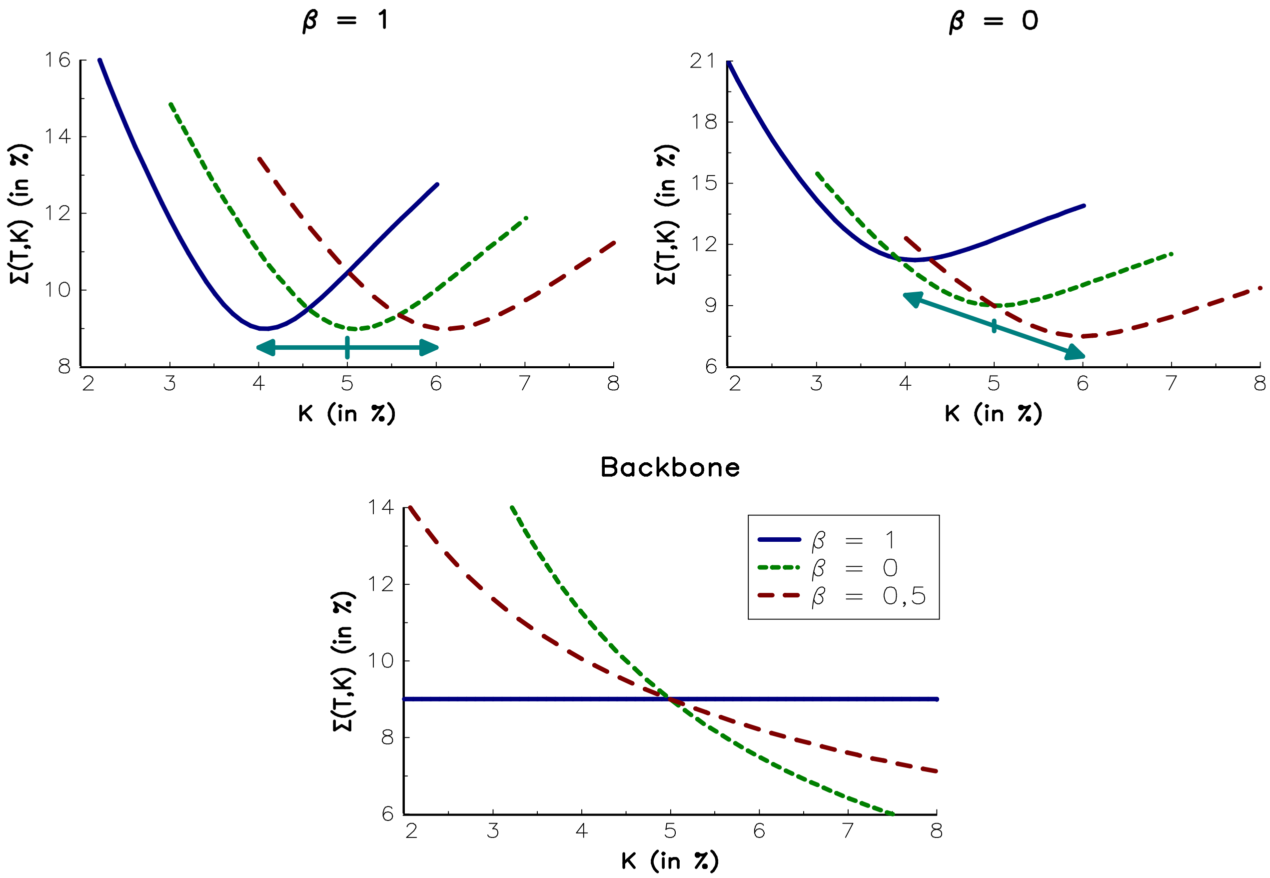


Figure: Impact of the parameter β

SABR model

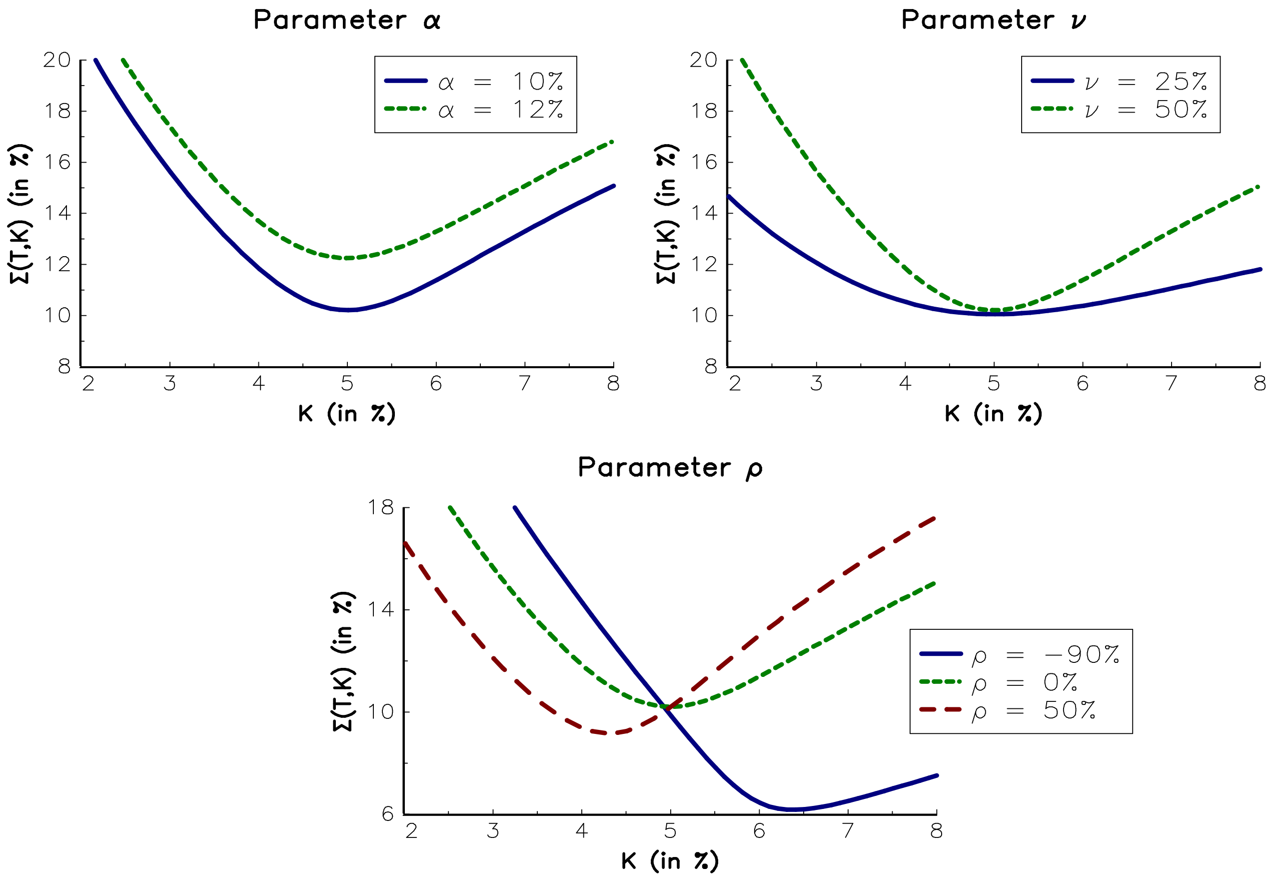


Figure: Impact of the parameters α, ν and ρ

SABR model

The parameters β and ρ impact the slope of the smile in a similar way. Then, they cannot be jointly identifiable. For example, let us consider the following smile when F_0 is equal to 5%: $\Sigma_B(1, 3\%) = 13\%$, $\Sigma_B(1, 4\%) = 10\%$, $\Sigma_B(1, 5\%) = 9\%$ and $\Sigma_B(1, 7\%) = 10\%$. If we calibrate this smile for different values of β, we obtain the following solutions:

β	α	ν	ρ
0.0	0.0044	0.3203	0.2106
0.5	0.0197	0.3244	0.0248
1.0	0.0878	0.3388	−0.1552

SABR model

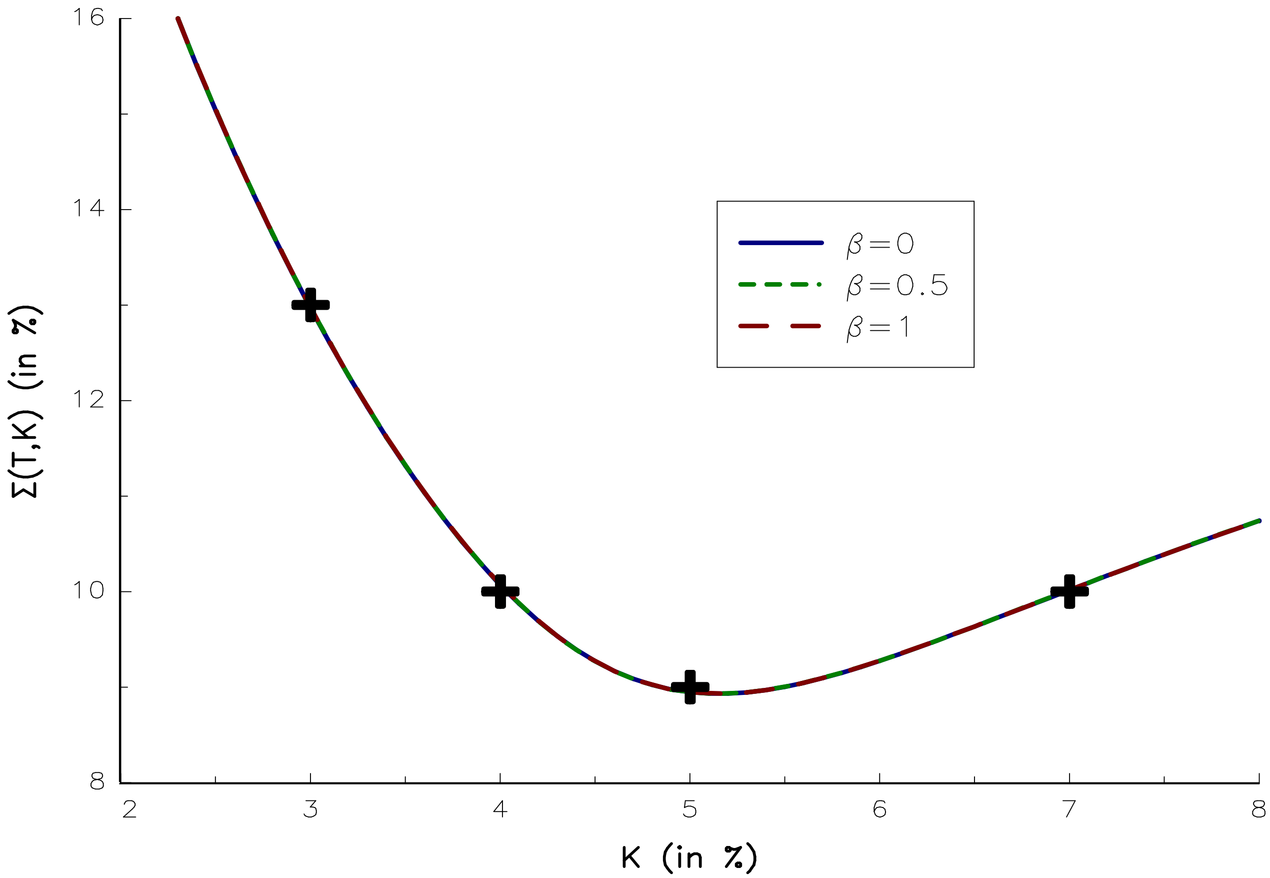


Figure: Implied volatility for different parameter sets (β, ρ)

SABR model

There are two approaches for estimating β:

1. β can be chosen from prior beliefs ($\beta = 0$ for the normal model, $\beta = 0.5$ for the CIR model and $\beta = 1$ for the log-normal model)
2. β can be statistically estimated by considering the dynamics of the forward rate. Indeed, we have

$$\Sigma_t \left(T, F_t \right) \simeq \frac{\alpha}{F_t^{1-\beta}}$$

We can consider the following linear regression:

$$\ln \Sigma_t \left(T, F_t \right) = \ln \alpha + \left(\beta - 1 \right) \ln F_t + u_t$$

SABR model

Table: Calibration of the parameter β in the SABR model

Rate	Level		Difference		Empirical quantile of $\hat{\beta}_{t,t+h}$				
	$\hat{\beta}$	R_c^2	$\hat{\beta}$	R_c^2	10%	25%	50%	75%	90%
1y1y	−0.06	0.91	0.59	0.15	−2.01	−0.14	0.71	1.00	2.17
1y5y	−0.29	0.87	0.32	0.27	−1.80	−0.28	0.73	1.11	2.76
1y10y	−0.37	0.80	0.34	0.22	−2.04	−0.23	0.71	1.11	2.69
5y1y	0.42	0.29	0.35	0.22	−1.58	−0.31	0.71	1.00	2.38
5y5y	−0.01	0.73	0.23	0.28	−2.12	−0.36	0.61	1.00	2.52
5y10y	−0.10	0.69	0.27	0.23	−1.99	−0.30	0.70	1.05	2.58
10y1y	0.96	0.00	0.28	0.20	−1.88	−0.20	0.80	1.07	2.43
10y5y	−0.10	0.65	0.28	0.20	−2.02	−0.29	0.73	1.02	2.76
10y10y	−0.47	0.73	0.27	0.20	−1.71	−0.24	0.85	1.07	2.93

SABR model

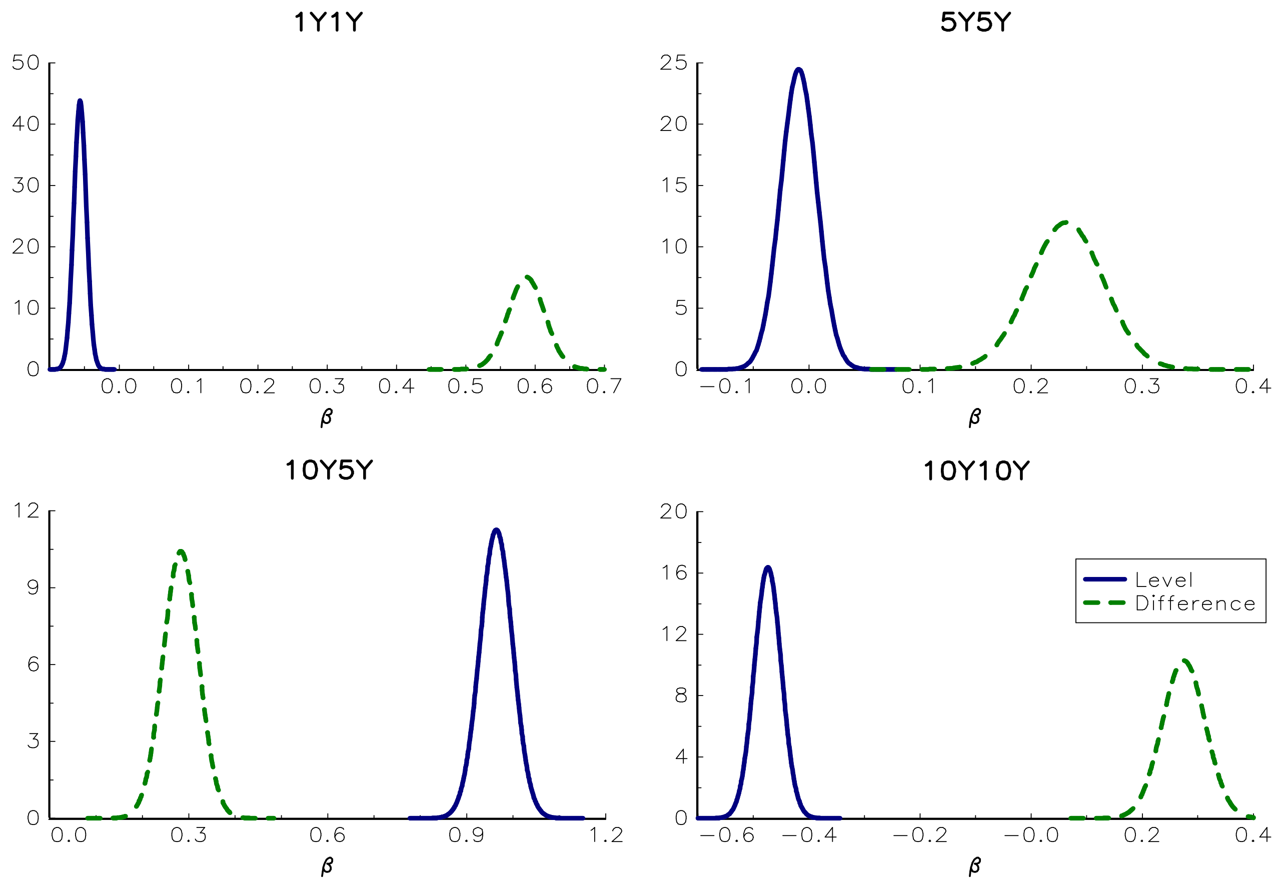


Figure: Probability density function of the estimate $\hat{\beta}$ (SABR model)

SABR model

- Once we have set the value of β, we estimate the parameters (α, ν, ρ) by fitting the observed implied volatilities
- However, we have seen that α is highly related to the ATM volatility. Indeed, we have:

$$\Sigma_B(T, F_0) = \frac{\alpha}{F_0^{1-\beta}} \left(1 + \left(\frac{(1-\beta)^2 \alpha^2}{24 F_0^{2-2\beta}} + \frac{\rho\alpha\nu\beta}{4 F_0^{1-\beta}} + \frac{2 - 3\rho^2}{24} \nu^2 \right) T \right)$$

- We deduce that:

$$\alpha^3 \left(\frac{(1-\beta)^2 T}{24 F_0^{2-2\beta}} \right) + \alpha^2 \left(\frac{\rho\nu\beta T}{4 F_0^{1-\beta}} \right) + \alpha \left(1 + \frac{2-3\rho^2}{24} \nu^2 T \right) - \Sigma_B(T, F_0) F_0^{1-\beta} = 0$$

- Let $\alpha = g_\alpha(\Sigma_B(T, F_0), \nu, \rho)$ be the positive root of the cubic equation. Therefore, imposing that the smile passes through the ATM volatility $\Sigma_B(T, F_0)$ allows to reduce the calibration to two parameters (ν, ρ)

SABR model

Example #10

We consider the following smile:

K (in %)	2.8	3.0	3.5	3.7	4.0	4.5	5.0	7.0
$\Sigma(T, K)$ (in %)	13.2	12.8	12.0	11.6	11.0	10.0	9.0	10.0

The maturity T is equal to one year and the forward rate F_0 is set to 5%

SABR model

If we consider a stochastic log-normal model ($\beta = 1$), we obtain the following results:

Calibration	α (in %)	β	ν	ρ (in %)	RSS	Σ_{ATM} (in %)
#1	9.466	1.00	0.279	-23.70	0.630	9.51
#2	8.944	1.00	0.322	-22.90	1.222	9.00

SABR model

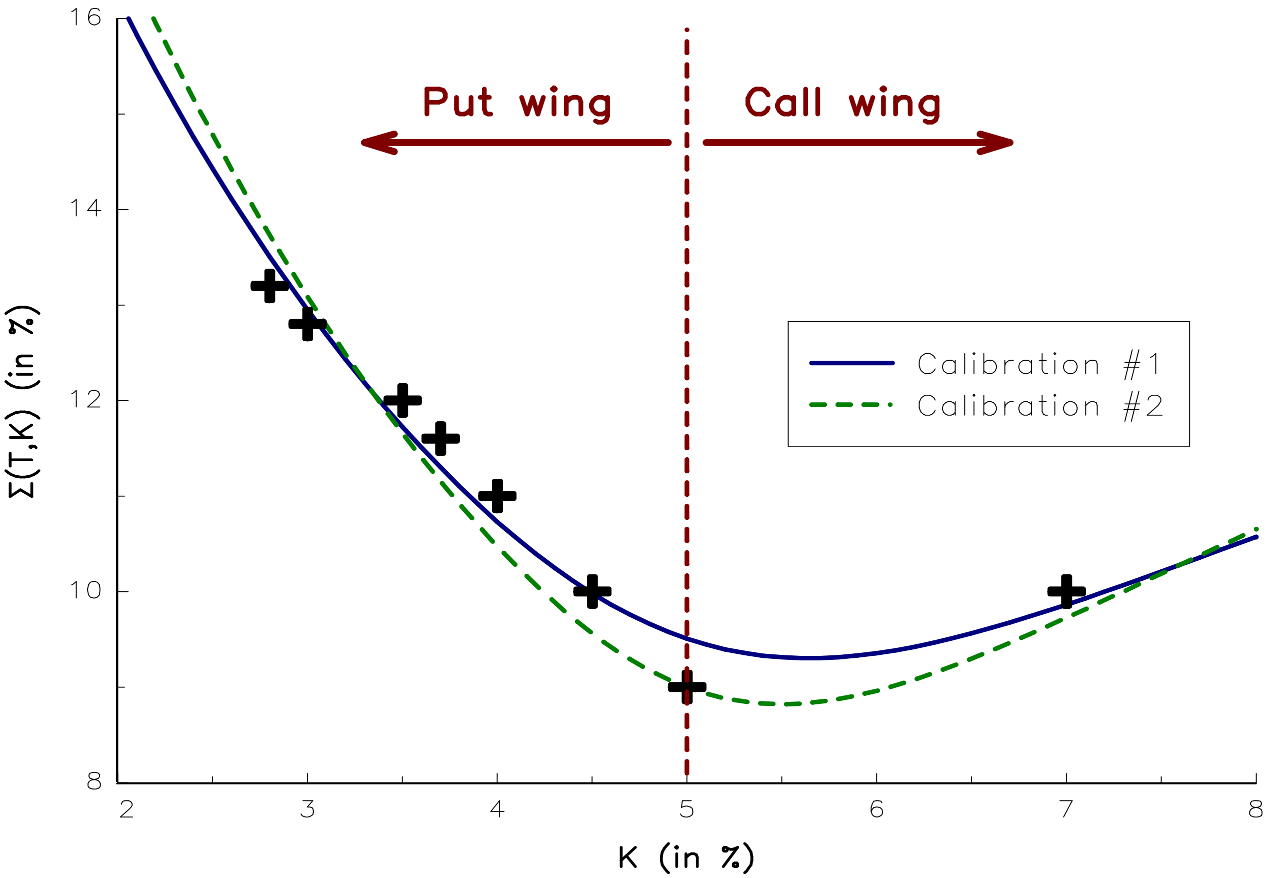


Figure: Calibration of the SABR model

SABR model

- The sensitivities correspond to the following formulas:

$$\boldsymbol{\Delta} = \frac{\partial \mathcal{C}_B}{\partial F_0} + \frac{\partial \mathcal{C}_B}{\partial \Sigma} \cdot \frac{\partial \Sigma_B(T, K)}{\partial F_0}$$

and:

$$\boldsymbol{\upsilon} = \frac{\partial \mathcal{C}_B}{\partial \Sigma} \cdot \frac{\partial \Sigma_B(T, K)}{\partial \alpha}$$

- To obtain these formulas, we apply the chain rule on the Black formula by assuming that the volatility Σ is not constant and depends on F_0 and α

SABR model

We notice that the vega is defined with respect to the parameter α. This approach is little used in practice, because it is difficult to hedge this model parameter. This is why traders prefer to compute the vega with respect to the ATM volatility:

$$\boldsymbol{\upsilon} = \frac{\partial\, \mathcal{C}_B}{\partial\, \Sigma} \cdot \frac{\partial\, \Sigma_B\left(T, K\right)}{\partial\, \alpha} \cdot \frac{\partial\, \alpha}{\partial\, \Sigma_{\mathrm{ATM}}}$$

where $\Sigma_{\mathrm{ATM}} = \Sigma_B\left(T, F_0\right)$

SABR model

Bartlett (2006) proposes a refinement for computing the delta. Indeed, a shift in F_0 produces a shift in α, because the two processes $F(t)$ and $\alpha(t)$ are correlated. Since we have:

$$\begin{aligned} \mathrm{d}\alpha(t) &= \nu\alpha(t)\,\mathrm{d}W_2^Q(t) \\ &= \nu\alpha(t)\left(\rho\,\mathrm{d}W_1^Q(t) + \sqrt{1-\rho^2}\,\mathrm{d}W(t)\right) \end{aligned}$$

and:

$$\mathrm{d}W_1^Q(t) = \frac{\mathrm{d}F(t)}{\alpha(t)\,F(t)^{\beta}}$$

we deduce that:

$$\mathrm{d}\alpha(t) = \frac{\nu\rho}{F(t)^{\beta}}\,\mathrm{d}F(t) + \nu\alpha(t)\sqrt{1-\rho^2}\,\mathrm{d}W(t)$$

SABR model

The new delta is then:

$$\begin{aligned}
\boldsymbol{\Delta}^{\star} &= \frac{\partial\,\mathcal{C}_B}{\partial\,F_0} + \frac{\partial\,\mathcal{C}_B}{\partial\,\Sigma}\left(\frac{\partial\,\Sigma_B\left(T,K\right)}{\partial\,F_0} + \frac{\partial\,\Sigma_B\left(T,K\right)}{\partial\,\alpha}\cdot\frac{\partial\,\alpha}{\partial\,F_0}\right) \\
&= \frac{\partial\,\mathcal{C}_B}{\partial\,F_0} + \frac{\partial\,\mathcal{C}_B}{\partial\,\Sigma}\left(\frac{\partial\,\Sigma_B\left(T,K\right)}{\partial\,F_0} + \frac{\nu\rho}{F\left(t\right)^{\beta}}\frac{\partial\,\Sigma_B\left(T,K\right)}{\partial\,\alpha}\right) \\
&= \boldsymbol{\Delta} + \frac{\nu\rho}{F\left(t\right)^{\beta}}\boldsymbol{\upsilon}
\end{aligned}$$

$\Rightarrow$ The new delta incorporates a part of the vega risk

Factor models

- Factor models: Vasicek, CIR, HJM, etc.
- Interest rates are linked to some factors $X(t)$, which can be observable or not observable
- The factor is directly the instantaneous interest rate $r(t)$ in Vasicek or CIR models
- Multi-factor models by considering explicit factors (level, slope, convexity, etc.)
- Professional practice based on non-explicit factors

Linear and quadratic Gaussian models

- Let us assume that the instantaneous interest rate $r(t)$ is linked to the factors $X(t)$ under the risk-neutral probability $\mathbb{Q}$ as follows:

$$r(t) = \alpha(t) + \beta(t)^\top X(t) + X(t)^\top \Gamma(t) X(t)$$

  where $\alpha(t)$ is a scalar, $\beta(t)$ is a $n \times 1$ vector and $\Gamma(t)$ is a $n \times n$ matrix

- The factors follow an Ornstein-Uhlenbeck process:

$$\mathrm{d}X(t) = (a(t) + B(t) X(t)) \, \mathrm{d}t + \Sigma(t) \, \mathrm{d}W^{\mathbb{Q}}(t)$$

  where $a(t)$ is a $n \times 1$ vector, $B(t)$ is a $n \times n$ matrix, $\Sigma(t)$ is a $n \times n$ matrix and $W^{\mathbb{Q}}(t)$ is a standard n-dimensional Brownian motion

Linear and quadratic Gaussian models

There exists a family of $\hat{\alpha}(t,T)$, $\hat{\beta}(t,T)$ and $\hat{\Gamma}(t,T)$ such that the price of the zero-coupon bond $B(t,T)$ is given by:

$$B(t,T) = \exp\left(-\hat{\alpha}(t,T) - \hat{\beta}(t,T)^\top X(t) - X(t)^\top \hat{\Gamma}(t,T) X(t)\right)$$

where $\hat{\alpha}(t,T)$, $\hat{\beta}(t,T)$ and $\hat{\Gamma}(t,T)$ solve a system of Riccati equations. If we assume that the matrix $\hat{\Gamma}(t,T)$ is symmetric, we obtain:

$$\begin{array}{rcl} \partial_t \hat{\alpha}(t,T) & = & -\operatorname{tr}\left(\Sigma(t)\Sigma(t)^\top \hat{\Gamma}(t,T)\right) - \hat{\beta}(t,T)^\top a(t) + \\ & & \frac{1}{2}\hat{\beta}(t,T)^\top \Sigma(t)\Sigma(t)^\top \hat{\beta}(t,T) - \alpha(t) \\ \partial_t \hat{\beta}(t,T) & = & -B(t)^\top \hat{\beta}(t,T) + 2\hat{\Gamma}(t,T)\Sigma(t)\Sigma(t)^\top \hat{\beta}(t,T) - \\ & & 2\hat{\Gamma}(t,T) a(t) - \beta(t) \\ \partial_t \hat{\Gamma}(t,T) & = & 2\hat{\Gamma}(t,T)\Sigma(t)\Sigma(t)^\top \hat{\Gamma}(t,T) - \\ & & 2\hat{\Gamma}(t,T) B(t) - \Gamma(t) \end{array}$$

with the boundary conditions $\hat{\alpha}(T,T) = \hat{\beta}(T,T) = \hat{\Gamma}(T,T) = \mathbf{0}$

Linear and quadratic Gaussian models

The forward interest rate $F(t, T_1, T_2)$ is given by:

$$\begin{aligned} F(t, T_1, T_2) &= -\frac{1}{T_2 - T_1} \ln \frac{B(t, T_2)}{B(t, T_1)} \\ &= \frac{\hat{\alpha}(t, T_2) - \hat{\alpha}(t, T_1) + \left(\hat{\beta}(t, T_2) - \hat{\beta}(t, T_2)\right)^\top X(t)}{T_2 - T_1} + \\ & \quad \frac{X(t)^\top \left(\hat{\Gamma}(t, T_2) - \hat{\Gamma}(t, T_1)\right) X(t)}{T_2 - T_1} \end{aligned}$$

We deduce that the instantaneous forward rate is equal to:

$$f(t, T) = \alpha(t, T) + \beta(t, T)^\top X(t) + X(t)^\top \Gamma(t, T) X(t)$$

where $\alpha(t, T) = \partial_T \hat{\alpha}(t, T)$, $\beta(t, T) = \partial_T \hat{\beta}(t, T)$ and $\Gamma(t, T) = \partial_T \hat{\Gamma}(t, T)$

It follows that $\alpha(t) = \alpha(t, t) = \partial_t \hat{\alpha}(t, t)$, $\beta(t) = \beta(t, t) = \partial_t \hat{\beta}(t, t)$ and $\Gamma(t) = \Gamma(t, t) = \partial_t \hat{\Gamma}(t, t)$

Linear and quadratic Gaussian models

Let $V(t, X)$ be the price of the option, whose payoff is $f(x)$. It satisfies the following PDE:

$$\begin{aligned}\frac{1}{2}\operatorname{trace}\left(\Sigma(t)\,\partial_X^2 V(t,X)\,\Sigma(t)^\top\right) + (a(t) + B(t)X)\,\partial_X V(t,X) + \\ \partial_t V(t,X) - \left(\alpha(t) + \beta(t)^\top X + X^\top \Gamma(t) X\right) V(t,X) &= 0\end{aligned}$$

Once we have specified the functions $\alpha(t)$, $\beta(t)$, $\Gamma(t)$, $a(t)$, $B(t)$ and $\Sigma(t)$, we can then price the option by solving numerically the previous multidimensional PDE with the terminal condition $V(T, X) = f(X)$ Most of the time, the payoff is not specified with respect to the state variables X, but depends on the interest rate $r(t)$. In this case, we use the following transformation:

$$f(r) = f\left(\alpha(T) + \beta(T)^\top X + X^\top \Gamma(T) X\right)$$

Dynamics of risk factors under the forward probability measure

We have:

$$\frac{\mathrm{d}B(t,T)}{B(t,T)} = r(t)\,\mathrm{d}t - \left(2\hat{\Gamma}(t,T)X(t) + \hat{\beta}(t,T)\right)^{\top}\Sigma(t)\,\mathrm{d}W^{\mathbb{Q}}(t)$$

We deduce that:

$$W^{\mathbb{Q}^{\star}(T)}(t) = W^{\mathbb{Q}}(t) + \int_0^t \Sigma(s)^{\top}\left(2\hat{\Gamma}(s,T)X(s) + \hat{\beta}(s,T)\right)\mathrm{d}s$$

defines a Brownian motion under $\mathbb{Q}^{\star}(T)$

Dynamics of risk factors under the forward probability measure

It follows that:

$$\mathrm{d}X(t) = \left(\tilde{a}(t) + \tilde{B}(t)X(t)\right)\mathrm{d}t + \Sigma(t)\,\mathrm{d}W^{\mathbb{Q}^\star(T)}(t)$$

where:

$$\tilde{a}(t) = a(t) - \Sigma(t)\Sigma(t)^\top \hat{\beta}(t,T)$$

and:

$$\tilde{B}(t) = B(t) - 2\Sigma(t)\Sigma(t)^\top \hat{\Gamma}(t,T)$$

We conclude that $X(t)$ is Gaussian under any forward probability measure $\mathbb{Q}^\star(T)$:

$$X(t) \sim \mathcal{N}(m(0,t), V(0,t))$$

Dynamics of risk factors under the forward probability measure

El Karoui *et al.* (1992) showed that the conditional mean and variance satisfies the following forward differential equations:

$$\partial_T m(t,T) = a(T) + B(T)\, m(t,T) - 2V(t,T)\,\Gamma(T)\, m(t,T) - V(t,T)\,\beta(T)$$

and:

$$\begin{array}{rcl}\partial_T V(t,T) & = & V(t,T)\, B(T)^\top + B(T)\, V(t,T) - 2V(t,T)\,\Gamma(T)\, V(t,T) + \\ & & \Sigma(T)\,\Sigma(T)^\top\end{array}$$

- If t is equal to zero, the initial conditions are $m(0,0) = X(0) = \mathbf{0}$ and $V(0,0) = \mathbf{0}$
- If $t \neq 0$, we proceed in two steps: first, we calculate numerically the solutions $m(0,t)$ and $V(0,t)$, and second, we initialize the system with $m(t,t) = m(0,t)$ and $V(t,t) = V(0,t)$

Dynamics of risk factors under the forward probability measure

In fact, the previous forward differential equations are not obtained under the traditional forward probability measure $\mathbb{Q}^{\star}(T)$, but under the probability measure $\mathbb{Q}^{\star}(t,T)$ defined by the following Radon-Nykodin derivative:

$$\frac{\mathrm{d}\mathbb{Q}^{\star}(t,T)}{\mathrm{d}\mathbb{P}} = e^{-\int_0^T r(s)\,\mathrm{d}s} e^{\int_t^T f(t,s)\,\mathrm{d}s}$$

The reason is that we would like to price at time t any caplet with maturity T. Therefore, this is the maturity T and not the filtration $\mathcal{F}_t$ that moves

Pricing caplets and swaptions

The formula of the Libor rate $L(t, T_{i-1}, T_i)$ at time t between the dates T_{i-1} and T_i is:

$$L(t, T_{i-1}, T_i) = \frac{1}{T_i - T_{i-1}} \left(\frac{B(t, T_{i-1})}{B(t, T_i)} - 1 \right)$$

It follows that the price of the caplet is given by:

$$\mathrm{Caplet} = B(0, t) \mathbb{E}^{\mathbb{Q}^\star(t)} \left[\left(B(t, T_{i-1}) - (1 + (T_i - T_{i-1}) K) B(t, T_i) \right)^+ \right]$$

where $\mathbb{Q}^\star(t)$ is the forward probability measure. We can then calculate the price using two approaches:

1. we can solve the partial differential equation
2. we can calculate the mathematical expectation using numerical integration

Pricing caplets and swaptions

In the first approach, we consider the PDE with the following payoff:

$$f(X) = \max(0, g(X))$$

where:

$$\begin{aligned} g(X) &= \exp\left(-\hat{\alpha}(t, T_{i-1}) - \hat{\beta}(t, T_{i-1})^\top X - X^\top \hat{\Gamma}(t, T_{i-1}) X\right) - \\ & \quad (1 + \delta_{i-1} K) \exp\left(-\hat{\alpha}(t, T_i) - \hat{\beta}(t, T_i)^\top X - X^\top \hat{\Gamma}(t, T_i) X\right) \end{aligned}$$

In the second approach, we have $X(t) \sim \mathcal{N}(m(0,t), V(0,t))$ under the forward probability $\mathbb{Q}^\star(t)$. We deduce that:

$$\text{Caplet}(t, T_{i-1}, T_i) = B(0,t) \int f(x) \phi_n(x; m(0,t), V(0,t)) \,\mathrm{d}x$$

This integral can be computed numerically using Gauss-Legendre quadrature methods

Calibration and practice of factor models

- The calibration of the model consists in fitting the functions $\alpha(t)$, $\beta(t)$, $\Gamma(t)$, $a(t)$, $B(t)$ and $\Sigma(t)$
- Generally, professionals assume that $a(t) = 0$ and $B(t) = \mathbf{0}$

Calibration and practice of factor models

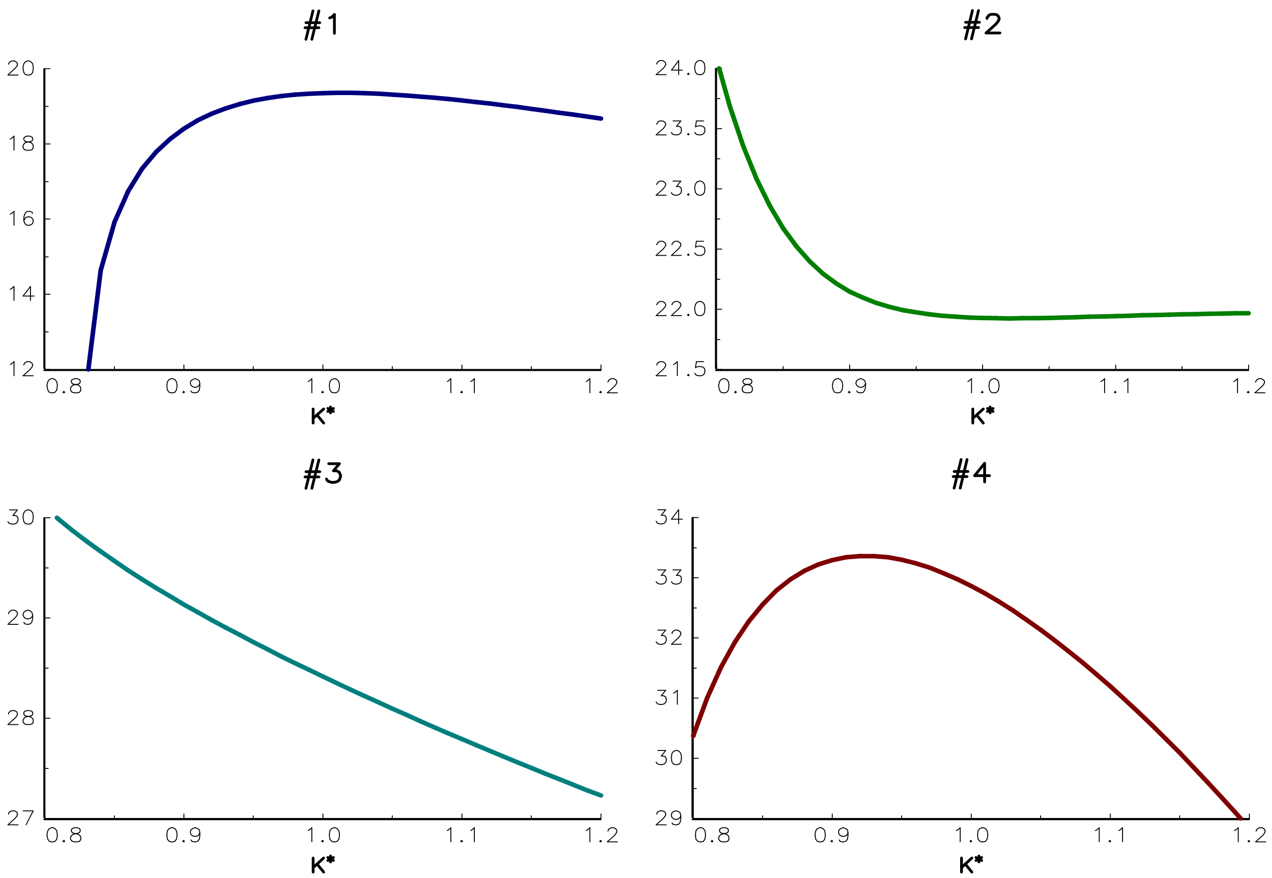


Figure: Volatility smiles generated by the quadratic Gaussian model

Impact of dividends on option prices

- Let us consider that the underlying asset pays a continuous dividend yield d during the life of the option
- The risk-neutral dynamics become:

$$\mathrm{d}S(t) = (r - d)\, S(t)\, \mathrm{d}t + \sigma S(t)\, \mathrm{d}W(t)$$

- We deduce that the Black-Scholes formula is equal to:

$$\mathcal{C}_0 = S_0 e^{-dT}\Phi(d_1) - K e^{-rT}\Phi(d_2)$$

where:

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T}}\left(\ln\frac{S_0}{K} + (r-d)\, T\right) + \frac{1}{2}\sigma\sqrt{T} \\ d_2 &= d_1 - \sigma\sqrt{T} \end{aligned}$$

Impact of dividends on option prices

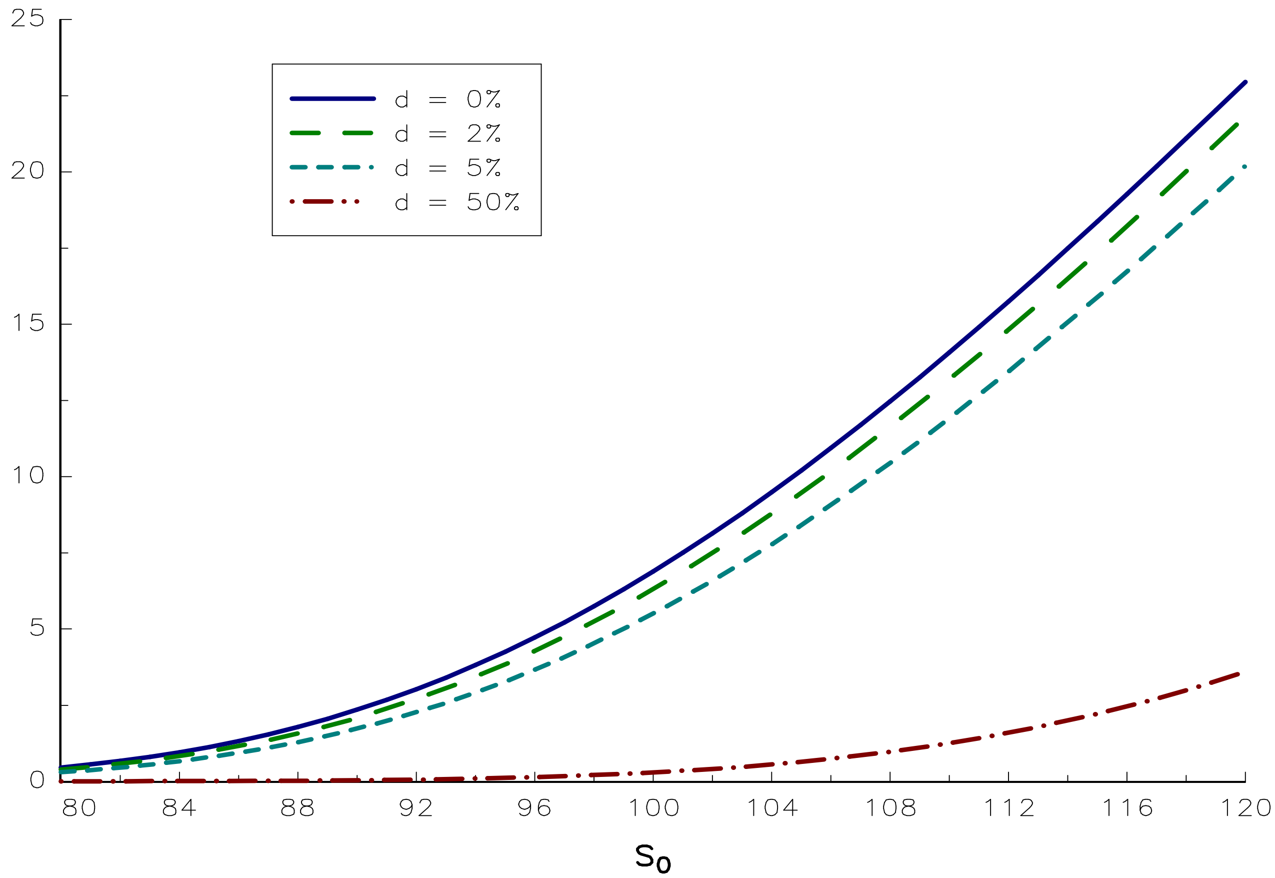


Figure: Impact of dividends on the call option price

Models of discrete dividends

We denote by $S(t)$ the market price and $Y(t)$ an additional process that is assumed to be a geometric Brownian motion:

$$\mathrm{d}Y(t) = rY(t)\,\mathrm{d}t + \sigma Y(t)\,\mathrm{d}W^{\mathbb{Q}}(t)$$

$\Rightarrow$ Three main approaches to take into account discrete dividends

Models of discrete dividends

1. $Y(t)$ is the capital price process excluding the dividends and the market price $S(t)$ is equal to the sum of the capital price and the discounted value of future dividends:

$$S(t) = Y(t) + \sum_{t_k \in [t,T]} D(t_k) e^{-r(t_k - t)}$$

To price European options, we then replace the price S_0 by the adjusted price $Y_0 = S_0 - \sum_{t_k \leq T} D(t_k) e^{-rt_k}$

Models of discrete dividends

2. We define $D(t)$ as the sum of capitalized dividends paid until time t:

$$D(t) = \sum \mathbb{1}\{t_k < t\} \cdot D(t_k) e^{r(t-t_k)}$$

The market price $S(t)$ is equal to the difference between the cum-dividend price $Y(t)$ and the capitalized dividends :

$$S(t) = Y(t) - D(t)$$

We deduce that:

$$\begin{aligned} (S(T) - K)^+ &= (Y(T) - D(T) - K)^+ \\ &= (Y(T) - (K + D(T)))^+ \\ &= (Y(T) - K')^+ \end{aligned}$$

In the case of European options, we replace the strike K by the adjusted strike $K' = K + \sum_{t_k \leq T} D(t_k) e^{r(T-t_k)}$

Models of discrete dividends

3. The last approach considers the market price process as a discontinuous process:

$$\left\{ \begin{array}{ll} \mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}W^{\mathbb{Q}}(t) & \text{if } t_{k-1} < t < t_k \\ S(t) = S\left(t_k^-\right) - D(t_k) & \text{if } t = t_k \end{array} \right.$$

Models of discrete dividends

Example #11

We assume that $S_0 = 100$, $K = 100$, $\sigma = 30\%$, $T = 1$, $r = 5\%$ and $b = 5\%$. A dividend $D(t_1)$ will be paid at time $t_1 = 0.5$

Table: Impact of the dividend on the option price

	Call			Put		
$D(t_1)$	(#1)	(#2)	(#3)	(#1)	(#2)	(#3)
0	14.23	14.23	14.23	9.35	9.35	9.35
3	12.46	12.81	12.69	10.51	10.86	10.64
5	11.34	11.92	11.69	11.34	11.92	11.59
10	8.78	9.93	9.42	13.66	14.80	14.20

The two-asset case

- We consider the example of a basket option on two assets
- Let $S_i(t)$ be the price process of asset i at time t. According to the Black-Scholes model, we have:

$$\left\{ \begin{array}{l} \mathrm{d}S_1(t) = b_1 S_1(t)\,\mathrm{d}t + \sigma_1 S_1(t)\,\mathrm{d}W_1^{\mathbb{Q}}(t) \\ \mathrm{d}S_2(t) = b_2 S_2(t)\,\mathrm{d}t + \sigma_2 S_2(t)\,\mathrm{d}W_2^{\mathbb{Q}}(t) \end{array} \right.$$

  where b_i and σ_i are the cost-of-carry and the volatility of asset i
- Under the risk-neutral probability measure $\mathbb{Q}$, $W_1^{\mathbb{Q}}(t)$ and $W_2^{\mathbb{Q}}(t)$ are two correlated Brownian motions:

$$\mathbb{E}\left[W_1^{\mathbb{Q}}(t)\,W_2^{\mathbb{Q}}(t)\right] = \rho\, t$$

The two-asset case

- The option price associated to the payoff $\left(\alpha_1 S_1(T) + \alpha_2 S_2(T) - K\right)^+$ is the solution of the two-dimensional PDE:

$$\begin{aligned} \frac{1}{2}\sigma_1^2 S_1^2 \partial_{S_1}^2 \mathcal{C} + \frac{1}{2}\sigma_2^2 S_2^2 \partial_{S_2}^2 \mathcal{C} + \rho\sigma_1\sigma_2 S_1 S_2 \partial_{S_1,S_2}^2 \mathcal{C} + & \\ b_1 S_1 \partial_{S_1} \mathcal{C} + b_2 S_2 \partial_{S_2} \mathcal{C} + \partial_t \mathcal{C} - r\mathcal{C} & = 0 \end{aligned}$$

  with the terminal condition:

$$\mathcal{C}(T, S_1, S_2) = \left(\alpha_1 S_1 + \alpha_2 S_2 - K\right)^+$$

- Using the Feynman-Kac representation theorem, we have:

$$\mathcal{C}_0 = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r\,\mathrm{d}t}\left(\alpha_1 S_1(T) + \alpha_2 S_2(T) - K\right)^+\right]$$

The two-asset case

- The value $\mathcal{C}_0$ can be calculated using numerical integration
- In some cases, the two-dimensional problem can be reduced to one-dimensional integration. For instance, if $\alpha_1 < 0$, $\alpha_2 > 0$ and $K > 0$, we obtain:

$$\mathcal{C}_0 = \int_{\mathbb{R}} \mathrm{BS}\left(S^{\star}(x), K^{\star}(x), \sigma^{\star}, T, b^{\star}, r\right) \phi(x) \, \mathrm{d}x$$

where:

$$\begin{aligned} S^{\star}(x) &= \alpha_2 S_2(0) e^{\rho \sigma_2 \sqrt{T} x} \\ K^{\star}(x) &= K - \alpha_1 S_1(0) e^{\left(b_1 - \frac{1}{2}\sigma_1^2\right) T + \sigma_1 \sqrt{T} x} \\ \sigma^{\star} &= \sigma_2 \sqrt{1 - \rho^2} \\ b^{\star} &= b_2 - \frac{1}{2} \rho^2 \sigma_2^2 \end{aligned}$$

The two-asset case

Example #12

We assume that $S_1(0) = S_2(0) = 100$, $\sigma_1 = \sigma_2 = 20\%$, $b_1 = 10\%$, $b_2 = 0$ and $r = 5\%$. We calculate the price of a basket option, whose maturity T is equal to one year. For the other characteristics (α_1, α_2, K), we consider different set of parameters: $(1, -1, 1)$, $(1, -1, 5)$, $(0.5, 0.5, 100)$, $(0.5, 0.5, 110)$ and $(0.1, 0.1, -5)$

The two-asset case

Table: Impact of the correlation on the basket option price

α_1		1.0	1.0	0.5	0.5	0.1
α_2		−1.0	−1.0	0.5	0.5	0.1
K		1	5	100	110	−5
ρ	−0.90	20.41	18.23	5.39	0.66	24.78
	−0.75	19.81	17.62	6.06	1.35	24.78
	−0.50	18.76	16.55	6.97	2.31	24.78
	−0.25	17.61	15.37	7.73	3.12	24.78
	0.00	16.35	14.08	8.39	3.83	24.78
	0.25	14.94	12.61	8.99	4.46	24.78
	0.50	13.30	10.88	9.54	5.05	24.78
	0.75	11.29	8.66	10.05	5.59	24.78
	0.90	9.78	6.81	10.34	5.90	24.78

Cega sensitivity

Table: Relationship between the basket option price and the correlation parameter ρ

Option type	Payoff	Increasing	Decreasing
Spread	$(S_2 - S_1 - K)^+$		✓
Basket	$(\alpha_1 S_1 + \alpha_2 S_2 - K)^+$	$\alpha_1\alpha_2 > 0$	$\alpha_1\alpha_2 < 0$
Max	$(\max(S_1, S_2) - K)^+$		✓
Min	$(\min(S_1, S_2) - K)^+$	✓	
Best-of call/call	$\max\left((S_1 - K_1)^+, (S_2 - K_2)^+\right)$		✓
Best-of put/put	$\max\left((K_1 - S_1)^+, (K_2 - S_2)^+\right)$		✓
Worst-of call/call	$\min\left((S_1 - K_1)^+, (S_2 - K_2)^+\right)$	✓	
Option!Worst-of Worst-of put/put	$\min\left((K_1 - S_1)^+, (K_2 - S_2)^+\right)$	✓	

Cega sensitivity

The sensitivity of the option price with respect to the correlation parameter ρ is called the cega:

$$\boldsymbol{c} = \frac{\partial\, \boldsymbol{\mathcal{C}}_0}{\partial \rho}$$

The previous analysis leads us to define the lower and upper bounds of the option price when the cega is either positive or negative:

$$\boldsymbol{\mathcal{C}}_0 \in \left\{ \begin{array}{ll} \left[\boldsymbol{\mathcal{C}}_0\left(\rho^-\right), \boldsymbol{\mathcal{C}}_0\left(\rho^+\right)\right] & \text{if } \boldsymbol{c} \geq 0 \\ \left[\boldsymbol{\mathcal{C}}_0\left(\rho^+\right), \boldsymbol{\mathcal{C}}_0\left(\rho^-\right)\right] & \text{if } \boldsymbol{c} \leq 0 \end{array} \right.$$

We can define the conservative price by taking the maximum between $\boldsymbol{\mathcal{C}}_0\left(\rho^-\right)$ and $\boldsymbol{\mathcal{C}}_0\left(\rho^+\right)$

Cega sensitivity

In the case where $\rho^- = -1$ and $\rho^+ = 1$, the bounds satisfy the one-dimensional PDE:

$$\left\{ \begin{array}{l} \frac{1}{2}\sigma_1^2 S^2 \partial_S^2 \mathcal{C}(t,S) + b_1 S \partial_S \mathcal{C}(t,S) + \partial_t \mathcal{C}(t,S) - r\mathcal{C}(t,S) = 0 \\ \mathcal{C}(T,S) = f(S, g(S)) \end{array} \right.$$

where:

$$g(S) = S_2(0) \left(\frac{S}{S_1(0)} \right)^{\pm\sigma_2/\sigma_1} \exp\left(\left(b_2 - \frac{1}{2}\sigma_2^2 \pm \left(\frac{1}{2}\sigma_1\sigma_2 - \frac{\sigma_2}{\sigma_1} b_1 \right) \right) T \right)$$

The implied correlation

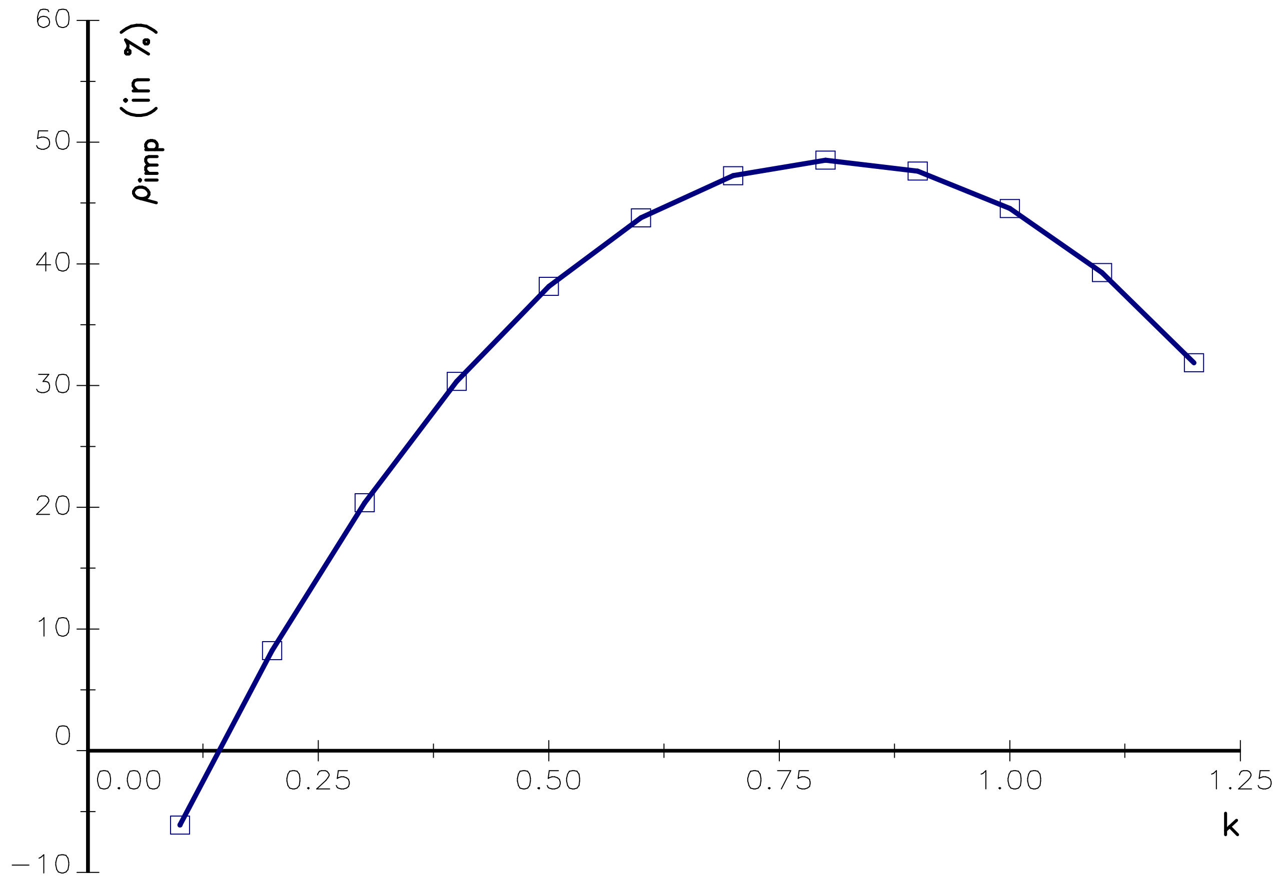


Figure: Correlation smile

Riding on the smiles

- In practice, the two volatilities are unknown and must be deduced from the volatility smiles $\Sigma_1(K_1, T)$ and $\Sigma_2(K_2, T)$ of the two assets
- The difficulty is then to find the corresponding strikes K_1 and K_2

Riding on the smiles

- In the case of the general payoff $(\alpha_1 S_1(T) + \alpha_2 S_2(T) - K)^+$, we have:

$$\begin{cases} (\alpha_1 = 1, \alpha_2 = 0, K \geq 0) \Rightarrow K_1 = K \\ (\alpha_1 = -1, \alpha_2 = 0, K \leq 0) \Rightarrow K_1 = -K \end{cases}$$

and:

$$\begin{cases} (\alpha_1 = 0, \alpha_2 = 1, K \geq 0) \Rightarrow K_2 = K \\ (\alpha_1 = 0, \alpha_2 = -1, K \leq 0) \Rightarrow K_2 = -K \end{cases}$$

- The payoff of the spread option can be written as follows:

$$\begin{aligned} (S_1(T) - S_2(T) - K)^+ &= ((S_1(T) - K_1) + (K_2 - S_2(T)))^+ \\ &\leq \underbrace{(S_1(T) - K_1)^+}_{\text{Call}} + \underbrace{(K_2 - S_2(T))^+}_{\text{Put}} \end{aligned}$$

where $K_1 = K_2 + K$

- Therefore, the price of the spread option can be bounded above by a call price on S_1 plus a put price on S_2
- However, the implicit strikes can take different values

Riding on the smiles

Let us assume that $S_1(0) = S_2(0) = 100$ and $K = 4$. Below, we give five pairs (K_1, K_2) and the associated implied volatilities $(\Sigma_1(K_1, T), \Sigma_2(K_2, T))$:

Pair	#1	#2	#3	#4	#5
K_1	104	103	102	101	100
K_2	100	99	98	97	96
$\Sigma_1(K_1, T)$	16%	17%	18%	19%	20%
$\Sigma_2(K_2, T)$	20%	22%	24%	26%	28%
$\mathcal{C}_0$	10.77	11.37	11.99	12.61	13.24

The multi-asset case

How to define a conservative price?

- In the multivariate case, the PDE becomes:

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^{n}\sigma_i^2 S_i^2 \partial_{S_i}^2 \mathcal{C} + \sum_{i<j}^{n} \rho_{i,j}\sigma_i\sigma_j S_i S_j \partial_{S_i,S_j}^2 \mathcal{C} + & \\ \sum_{i=1}^{n} b_i S_i \partial_i \mathcal{C} + \partial_t \mathcal{C} - r\mathcal{C} &= 0\end{aligned}$$

  with the terminal value:

$$\mathcal{C}(T, S_1, \ldots, S_n) = f(S_1(T), \ldots, S_n(T))$$

- Here, $\rho_{i,j}$ is the correlation between the Brownian motions of S_i and S_j
- Most of the time, the trader uses the same value ρ for all asset correlations $\rho_{i,j}$

The multi-asset case

How to define a conservative price?

- We can show that the price is increasing (resp. decreasing) with respect to ρ if $\sum_{i<j}^{n} \sigma_i \sigma_j \partial^2_{S_i,S_j} f$ is a positive (resp. negative) measure
- Let us consider the payoff function $f(S_1, S_{2,} S_3) = (S_1 + S_2 - S_3 - K)^+$, we have:

$$\sum_{i<j}^{n} \sigma_i \sigma_j \partial^2_{S_i,S_j} f = (\sigma_1 \sigma_2 - \sigma_1 \sigma_3 - \sigma_2 \sigma_3) \cdot \mathbb{1} \{S_1 + S_2 - S_3 - K = 0\}$$

- If $\sigma_1 \sigma_2 - \sigma_1 \sigma_3 - \sigma_2 \sigma_3 > 0$, the price increases with respect to ρ, and if $\sigma_1 \sigma_2 - \sigma_1 \sigma_3 - \sigma_2 \sigma_3 < 0$, the price decreases with respect to ρ

The multi-asset case

Issues with constant correlation matrices

- We consider a basket of n stocks
- The basket volatility is given by:

$$\sigma_B = \sqrt{\sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2 \sum_{i>j}^{n} \rho_{i,j} w_i w_j \sigma_i \sigma_j}$$

where w_i is the weight of asset i in the basket, σ_i the volatility of asset i and $\rho_{i,j}$ the correlation between asset i and asset j

The multi-asset case

Issues with constant correlation matrices

- The implied correlation ρ_{imp} of the basket is defined as the root of the following equation:

$$\sigma_B^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2 - 2\rho_{\mathrm{imp}} \sum_{i>j}^{n} w_i w_j \sigma_i \sigma_j = 0$$

- We deduce that:

$$\rho_{\mathrm{imp}} = \frac{\sigma_B^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2}{2 \sum_{i>j}^{n} w_i w_j \sigma_i \sigma_j}$$

- Another expression of the implied correlation is:

$$\rho_{\mathrm{imp}} = \frac{\sigma_B^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2}{\left(\sum_{i=1}^{n} w_i \sigma_i\right)^2 - \sum_{i=1}^{n} w_i^2 \sigma_i^2}$$

The multi-asset case

Issues with constant correlation matrices

We consider the following payoff:

$$\left(S_1(T) - S_2(T) + S_3(T) - S_4(T) - K\right)_+ \cdot \mathbb{1}\left\{S_5(T) > L\right\}$$

We calculate the option price of maturity 3 months using the Black-Scholes model. We assume that $S_i(0) = 100$ and $\Sigma_i = 20\%$ for the five underlying assets, the strike K is equal to 5, the barrier L is equal to 105, and the interest rate r is set to 5%.

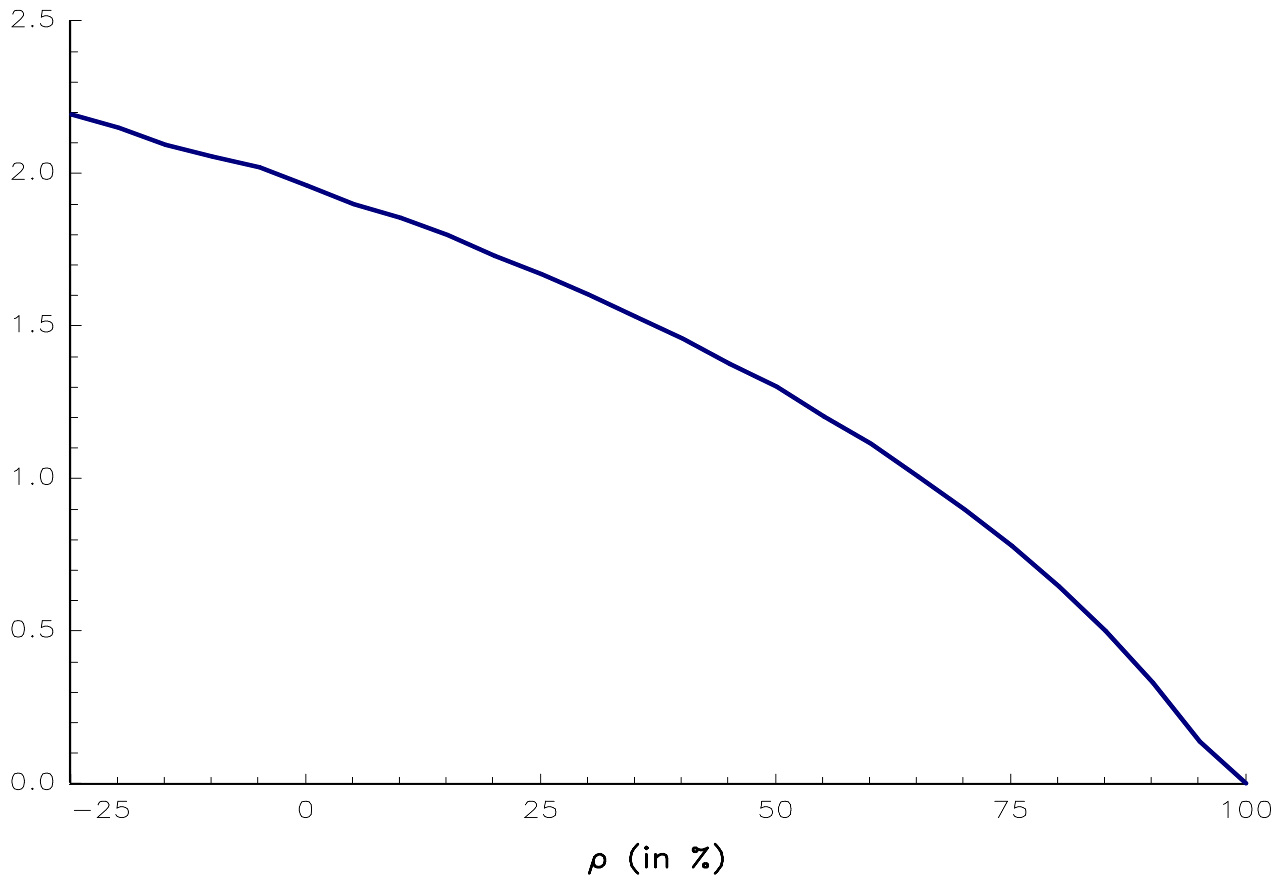
When the correlation matrix is $\mathbb{C}_5(\rho)$, the maximum price of 2.20 is not a conservative price. For instance, if we consider the correlation matrix below, we obtain an option price of 3.99:

$$\mathbb{C} = \begin{pmatrix} 1.0000 & 0.2397 & 0.7435 & -0.1207 & 0.0563 \\ 0.2397 & 1.0000 & -0.0476 & -0.0260 & -0.1958 \\ 0.7435 & -0.0476 & 1.0000 & 0.2597 & 0.1153 \\ -0.1207 & -0.0260 & 0.2597 & 1.0000 & -0.7568 \\ 0.0563 & -0.1958 & 0.1153 & -0.7568 & 1.0000 \end{pmatrix}$$

The multi-asset case

Issues with constant correlation matrices

Figure: Price of the basket option with respect to the constant correlation



Liquidity risk

Liquidity risk impacts trading costs of the hedging strategy

An example is the put option $\Rightarrow$ short strategy (can we be short on the underlying asset?)

Liquidity risk

Let us consider the replication of a call option. If the price of the underlying asset decreases sharply, the delta is reduced and the option trader has to sell asset shares. Because of their trend-following aspect, option traders generally buy assets when the market goes up and sell assets when the market goes down. However, we know that liquidity is asymmetric between these two market regimes. Therefore, it is more difficult to adjust the delta exposure when the market goes down, because of the lack of liquidity

Liquidity risk

Let us consider one of the most famous examples, which concerns call options on Sharpe ratio. Starting from 2004, some banks proposed to investors a payoff of the form $(\mathrm{SR}(0;T) - K)^+$ where $\mathrm{SR}(0;T)$ is the Sharpe ratio of the underlying asset during the option period. This payoff is relatively easy to replicate. However, most of call options on Sharpe ratio have been written on mutual funds and hedge funds. The difficulty comes from the liquidity of these underlying assets. For instance, the trader does not know exactly the price of the asset when he executes his order because of the notice period. This can be a big issue when the fund offers weekly or monthly liquidity. The second problem comes from the fact that the fund manager can impose lock-up period and gates. For instance, a gate limits the amount of withdrawals. During the 2008/2009 hedge fund crisis, many traders faced gate provisions and were unable to adjust their delta. This crisis marketed the end of call options on Sharpe ratio.

Exercises

- Option pricing models
  - Exercise 9.4.1 – Option pricing and martingale measure
  - Exercise 9.4.2 – The Vasicek model
  - Exercise 2.4.4 – Change of numéraire and Girsanov theorem
- Volatility
  - Exercise 9.4.8 – Dupire local volatility model
  - Exercise 2.4.9 – The stochastic normal model

References

Brigo, D., and Mercurio, F. (2006)
Interest Rate Models — Theory and Practice, Second edition, Springer.

Derman, E. (2001)
Markets and Models, *Risk Magazine*, 14(7), pp. 48-50.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 9.

Course 2023-2024 in Financial Risk Management
Lecture 9. Copulas and Extreme Value Theory

Thierry Roncalli*

*Amundi Asset Management[18]

*University of Paris-Saclay

September 2023

[18]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- **Lecture 9: Copulas and Extreme Value Theory**
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Sklar's theorem

A bi-dimensional copula is a function $\mathbf{C}$ which satisfies the following properties:

1. $\operatorname{Dom} \mathbf{C} = [0,1] \times [0,1]$
2. $\mathbf{C}(0,u) = \mathbf{C}(u,0) = 0$ and $\mathbf{C}(1,u) = \mathbf{C}(u,1) = u$ for all u in $[0,1]$
3. $\mathbf{C}$ is 2-increasing:

$$\mathbf{C}(v_1,v_2) - \mathbf{C}(v_1,u_2) - \mathbf{C}(u_1,v_2) + \mathbf{C}(u_1,u_2) \geq 0$$

for all $(u_1,u_2) \in [0,1]^2$, $(v_1,v_2) \in [0,1]^2$ such that $0 \leq u_1 \leq v_1 \leq 1$ and $0 \leq u_2 \leq v_2 \leq 1$

Remark

This means that $\mathbf{C}$ is a cumulative distribution function with uniform marginals:

$$\mathbf{C}(u_1,u_2) = \Pr\{U_1 \leq u_1, U_2 \leq u_2\}$$

where U_1 and U_2 are two uniform random variables

Sklar's theorem

We consider the function $\mathbf{C}^{\perp}(u_1, u_2) = u_1 u_2$. We have:

- $\mathbf{C}^{\perp}(0, u) = \mathbf{C}^{\perp}(u, 0) = 0$
- $\mathbf{C}^{\perp}(1, u) = \mathbf{C}^{\perp}(u, 1) = u$
- Since we have $v_2 - u_2 \geq 0$ and $v_1 \geq u_1$, it follows that $v_1(v_2 - u_2) \geq u_1(v_2 - u_2)$ and :

$$v_1 v_2 + u_1 u_2 - u_1 v_2 - v_1 u_2 \geq 0$$

$\Rightarrow$ $\mathbf{C}^{\perp}$ is a copula function and is called the product copula

Multivariate probability distribution with given marginals

Let $\mathbf{F}_1$ and $\mathbf{F}_2$ be two univariate distributions.
$\mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$ is a probability distribution with marginals $\mathbf{F}_1$ and $\mathbf{F}_2$:

- $u_i = \mathbf{F}_i(x_i)$ defines a uniform transformation ($u_i \in [0, 1]$)
- $\mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(\infty)) = \mathbf{C}(\mathbf{F}_1(x_1), 1) = \mathbf{F}_1(x_1)$

Sklar also shows that:

- Any bivariate distribution $\mathbf{F}$ admits a copula representation:

$$\mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$$

- The copula $\mathbf{C}$ is unique if the marginals are continuous

Multivariate probability distribution with given marginals

The Gumbel logistic distribution is equal to:

$$\mathbf{F}(x_1, x_2) = \left(1 + e^{-x_1} + e^{-x_2}\right)^{-1}$$

We have:

$$\mathbf{F}_1(x_1) \equiv \mathbf{F}(x_1, \infty) = \left(1 + e^{-x_1}\right)^{-1}$$

and $\mathbf{F}_2(x_2) \equiv (1 + e^{-x_2})^{-1}$. The quantile functions are then:

$$\mathbf{F}_1^{-1}(u_1) = \ln u_1 - \ln(1 - u_1)$$

and $\mathbf{F}_2^{-1}(u_2) = \ln u_2 - \ln(1 - u_2)$. We finally deduce that:

$$\mathbf{C}(u_1, u_2) = \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right) = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

is the Gumbel logistic copula

Expression of the copula density function

If the joint distribution function $\mathbf{F}(x_1, x_2)$ is absolutely continuous, we obtain:

$$\begin{aligned} f(x_1, x_2) &= \partial_{1,2}\, \mathbf{F}(x_1, x_2) \\ &= \partial_{1,2}\, \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \\ &= c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \end{aligned}$$

where $f(x_1, x_2)$ is the joint probability density function, f_1 and f_2 are the marginal densities and c is the copula density:

$$c(u_1, u_2) = \partial_{1,2}\, \mathbf{C}(u_1, u_2)$$

Remark

The condition $\mathbf{C}(v_1, v_2) - \mathbf{C}(v_1, u_2) - \mathbf{C}(u_1, v_2) + \mathbf{C}(u_1, u_2) \geq 0$ *is equivalent to* $\partial_{1,2}\, \mathbf{C}(u_1, u_2) \geq 0$ *when the copula density exists.*

Expression of the copula density function

In the case of the Gumbel logistic copula, we have:

$$\mathbf{C}(u_1, u_2) = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

and:

$$c(u_1, u_2) = \frac{2 u_1 u_2}{(u_1 + u_2 - u_1 u_2)^3}$$

Expression of the copula density function

We deduce that:

$$c(u_1, u_2) = \frac{f\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)}{f_1\left(\mathbf{F}_1^{-1}(u_1)\right) \cdot f_2\left(\mathbf{F}_2^{-1}(u_2)\right)}$$

If we consider the Normal copula, we have:

$$\mathbf{C}(u_1, u_2; \rho) = \Phi\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho\right)$$

and:

$$\begin{aligned} c(u_1, u_2; \rho) &= \frac{2\pi\left(1-\rho^2\right)^{-1/2} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)\right)}{(2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_1^2\right) \cdot (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_2^2\right)} \\ &= \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\frac{\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)}{(1-\rho^2)} + \frac{1}{2}\left(x_1^2 + x_2^2\right)\right) \end{aligned}$$

where $x_1 = \Phi_1^{-1}(u_1)$ and $x_2 = \Phi_2^{-1}(u_2)$

Expression of the copula density function

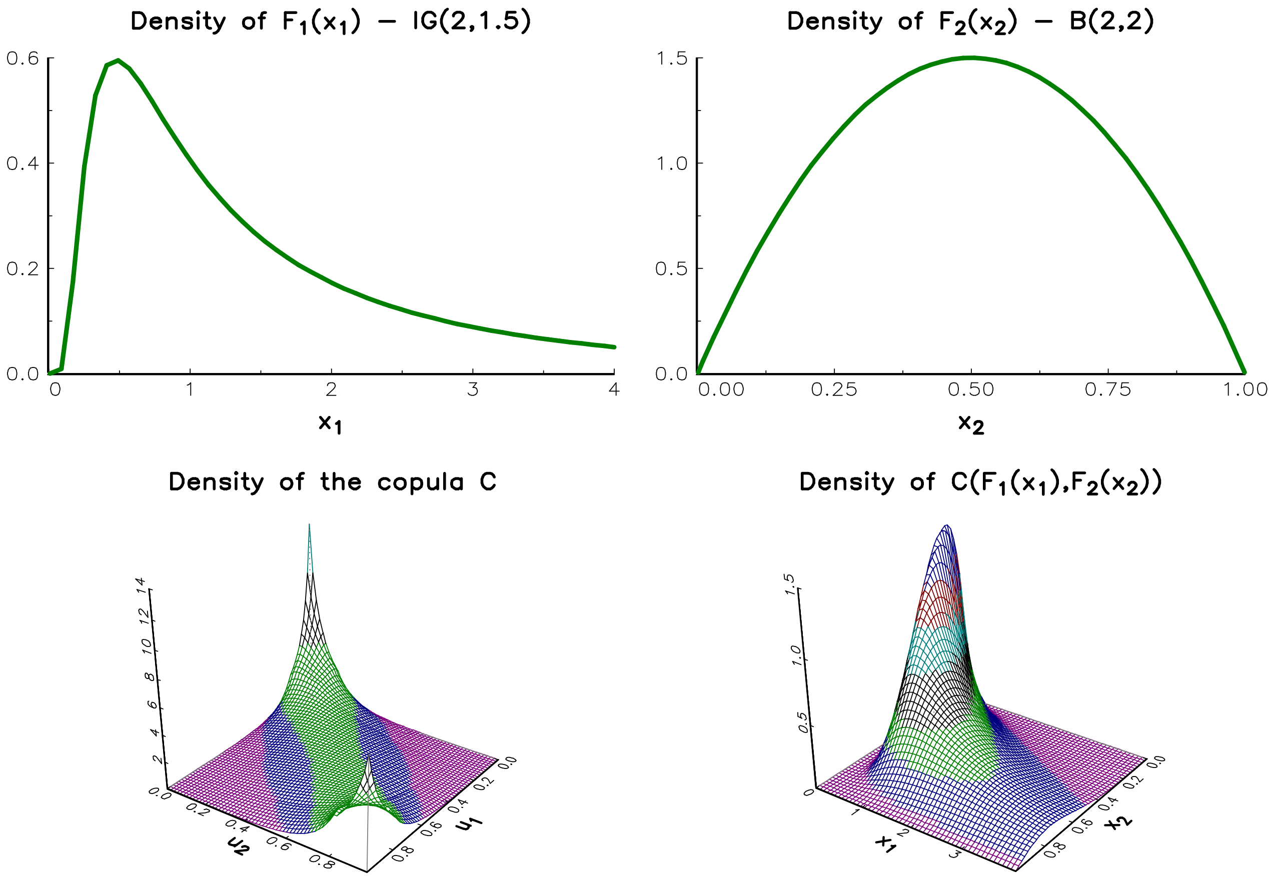


Figure: Construction of a bivariate probability distribution with given marginals and the Normal copula

Concordance ordering

Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two copula functions. We say that the copula $\mathbf{C}_1$ is smaller than the copula $\mathbf{C}_2$ and we note $\mathbf{C}_1 \prec \mathbf{C}_2$ if we have:

$$\mathbf{C}_1 (u_1, u_2) \leq \mathbf{C}_2 (u_1, u_2)$$

for all $(u_1, u_2) \in [0, 1]^2$

Let $\mathbf{C}_\theta (u_1, u_2) = \mathbf{C} (u_1, u_2; \theta)$ be a family of copula functions that depends on the parameter θ. The copula family $\{\mathbf{C}_\theta\}$ is totally ordered if, for all $\theta_2 \geq \theta_1$, $\mathbf{C}_{\theta_2} \succ \mathbf{C}_{\theta_1}$ (positively ordered) or $\mathbf{C}_{\theta_2} \prec \mathbf{C}_{\theta_1}$ (negatively ordered)

Remark

The Normal copula family is positively ordered

Fréchet bounds

We have:

$$\mathbf{C}^{-} \prec \mathbf{C} \prec \mathbf{C}^{+}$$

where:

$$\mathbf{C}^{-}(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$$

and:

$$\mathbf{C}^{+}(u_1, u_2) = \min(u_1, u_2)$$

The multivariate case

The canonical decomposition of a multivariate distribution function is:

$$\mathbf{F}(x_1, \ldots, x_n) = \mathbf{C}(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_n(x_n))$$

We have:

$$\mathbf{C}^- \prec \mathbf{C} \prec \mathbf{C}^+$$

where:

$$\mathbf{C}^-(u_1, \ldots, u_n) = \max\left(\sum_{i=1}^{n} u_i - n + 1, 0\right)$$

and:

$$\mathbf{C}^+(u_1, \ldots, u_n) = \min(u_1, \ldots, u_n)$$

Remark

$\mathbf{C}^-$ is not a copula when $n \geq 3$

Countermonotonicity and comonotonicity

Let $X = (X_1, X_2)$ be a random vector with distribution $\mathbf{F}$. We define the copula of (X_1, X_2) by the copula of $\mathbf{F}$:

$$\mathbf{F}(x_1, x_2) = \mathbf{C} \langle X_1, X_2 \rangle (\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$$

Definition

- X_1 and X_2 are countermonotonic – or $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^-$ – if there exists a random variable X such that $X_1 = f_1(X)$ and $X_2 = f_2(X)$ where f_1 and f_2 are respectively decreasing and increasing functions. In this case, $X_2 = f(X_1)$ where $f = f_2 \circ f_1^{-1}$ is a decreasing function
- X_1 and X_2 are independent if the dependence function is the product copula $\mathbf{C}^\perp$
- X_1 are X_2 are comonotonic – or $\mathbf{C} \langle X_1, X_2 \rangle = \mathbf{C}^+$ – if there exists a random variable X such that $X_1 = f_1(X)$ and $X_2 = f_2(X)$ where f_1 and f_2 are both increasing functions. In this case, $X_2 = f(X_1)$ where $f = f_2 \circ f_1^{-1}$ is an increasing function

Countermonotonicity and comonotonicity

- We consider a uniform random vector (U_1, U_2):

$$\begin{aligned}\mathbf{C}\langle U_1, U_2\rangle &= \mathbf{C}^- \Leftrightarrow U_2 = 1 - U_1 \\ \mathbf{C}\langle U_1, U_2\rangle &= \mathbf{C}^+ \Leftrightarrow U_2 = U_1\end{aligned}$$

- We consider a standardized Gaussian random vector (X_1, X_2). We have $U_1 = \Phi(X_1)$ and $U_2 = \Phi(X_2)$. We deduce that:

$$\begin{aligned}\mathbf{C}\langle X_1, X_2\rangle &= \mathbf{C}^- \Leftrightarrow \Phi(X_2) = 1 - \Phi(X_1) \Leftrightarrow X_2 = -X_1 \\ \mathbf{C}\langle X_1, X_2\rangle &= \mathbf{C}^+ \Leftrightarrow \Phi(X_2) = \Phi(X_1) \Leftrightarrow X_2 = X_1\end{aligned}$$

Countermonotonicity and comonotonicity

- We consider a random vector (X_1, X_2) where $X_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$. We have

$$U_i = \Phi\left(\frac{X_i - \mu_i}{\sigma_i}\right)$$

We deduce that:

$$\begin{aligned}
\mathbf{C}\left\langle X_1, X_2\right\rangle = \mathbf{C}^- \quad & \Leftrightarrow \quad \Phi\left(\frac{X_2 - \mu_2}{\sigma_2}\right) = 1 - \Phi\left(\frac{X_1 - \mu_1}{\sigma_1}\right) \\
& \Leftrightarrow \quad \Phi\left(\frac{X_2 - \mu_2}{\sigma_2}\right) = \Phi\left(-\frac{X_1 - \mu_1}{\sigma_1}\right) \\
& \Leftrightarrow \quad X_2 = \left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right) - \frac{\sigma_2}{\sigma_1}X_1
\end{aligned}$$

and:

$$\mathbf{C}\left\langle X_1, X_2\right\rangle = \mathbf{C}^+ \Leftrightarrow X_2 = \left(\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1\right) + \frac{\sigma_2}{\sigma_1}X_1$$

Countermonotonicity and comonotonicity

- We consider a random vector (X_1, X_2) where $X_i \sim \mathcal{LN}\left(\mu_i, \sigma_i^2\right)$. We have:

$$U_i = \Phi\left(\frac{\ln X_i - \mu_i}{\sigma_i}\right)$$

We deduce that:

$$\begin{aligned}
\mathbf{C}\langle X_1, X_2\rangle = \mathbf{C}^- \quad & \Leftrightarrow \quad \ln X_2 = \left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right) - \frac{\sigma_2}{\sigma_1}\ln X_1 \\
& \Leftrightarrow \quad X_2 = e^{\left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right)} e^{-\frac{\sigma_2}{\sigma_1}\ln X_1} \\
& \Leftrightarrow \quad X_2 = e^{\left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right)} X_1^{-\frac{\sigma_2}{\sigma_1}}
\end{aligned}$$

and:

$$\begin{aligned}
\mathbf{C}\langle X_1, X_2\rangle = \mathbf{C}^+ \quad & \Leftrightarrow \quad \ln X_2 = \left(\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1\right) + \frac{\sigma_2}{\sigma_1}\ln X_1 \\
& \Leftrightarrow \quad X_2 = e^{\left(\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1\right)} X_1^{\frac{\sigma_2}{\sigma_1}}
\end{aligned}$$

Countermonotonicity and comonotonicity

- If $X_1 \sim \mathcal{LN}(0,1)$ and $X_2 \sim \mathcal{LN}(0,1)$, we have:

$$\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^- \Leftrightarrow X_2 = \frac{1}{X_1}$$

- If $X_1 \sim \mathcal{LN}(0,2^2)$ and $X_2 \sim \mathcal{LN}(0,1)$, we have:

$$\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^+ \Leftrightarrow X_2 = \sqrt{X_1}$$

Linear dependence vs non-linear dependence

The concepts of counter- and comonotonicity concepts generalize the cases where the linear correlation of a Gaussian vector is equal to -1 or $+1$

Non-linear stochastic dependence

Scale invariance property

If h_1 and h_2 are two increasing functions on $\operatorname{Im} X_1$ and $\operatorname{Im} X_2$, then we have:

$$\mathbf{C}\left\langle h_1\left(X_1\right), h_2\left(X_2\right)\right\rangle = \mathbf{C}\left\langle X_1, X_2\right\rangle$$

Non-linear stochastic dependence

Proof (marginals)

We note $\mathbf{F}$ and $\mathbf{G}$ the probability distributions of the random vectors (X_1, X_2) and $(Y_1, Y_2) = (h_1(X_1), h_2(X_2))$. The marginals of $\mathbf{G}$ are:

$$\begin{aligned}
\mathbf{G}_1(y_1) &= \Pr\{Y_1 \leq y_1\} \\
&= \Pr\{h_1(X_1) \leq y_1\} \\
&= \Pr\left\{X_1 \leq h_1^{-1}(y_1)\right\} \quad \text{(because } h_1 \text{ is strictly increasing)} \\
&= \mathbf{F}_1\left(h_1^{-1}(y_1)\right)
\end{aligned}$$

and $\mathbf{G}_2(y_2) = \mathbf{F}_2\left(h_2^{-1}(y_2)\right)$. We deduce that $\mathbf{G}_1^{-1}(u_1) = h_1\left(\mathbf{F}_1^{-1}(u_1)\right)$ and $\mathbf{G}_2^{-1}(u_2) = h_2\left(\mathbf{F}_2^{-1}(u_2)\right)$

Non-linear stochastic dependence

Proof (copula)

By definition, we have:

$$\mathbf{C}\langle Y_1, Y_2\rangle(u_1, u_2) = \mathbf{G}\left(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)\right)$$

Moreover, it follows that:

$$\begin{aligned}
\mathbf{G}\left(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)\right) &= \Pr\left\{Y_1 \leq \mathbf{G}_1^{-1}(u_1), Y_2 \leq \mathbf{G}_2^{-1}(u_2)\right\} \\
&= \Pr\left\{h_1(X_1) \leq \mathbf{G}_1^{-1}(u_1), h_2(X_2) \leq \mathbf{G}_2^{-1}(u_2)\right\} \\
&= \Pr\left\{X_1 \leq h_1^{-1}\left(\mathbf{G}_1^{-1}(u_1)\right), X_2 \leq h_2^{-1}\left(\mathbf{G}_2^{-1}(u_2)\right)\right\} \\
&= \Pr\left\{X_1 \leq \mathbf{F}_1^{-1}(u_1), X_2 \leq \mathbf{F}_2^{-1}(u_2)\right\} \\
&= \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)
\end{aligned}$$

Because we have $\mathbf{C}\langle X_1, X_2\rangle(u_1, u_2) = \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)$, we deduce that:

$$\mathbf{C}\langle Y_1, Y_2\rangle = \mathbf{C}\langle X_1, X_2\rangle$$

Non-linear stochastic dependence

We have:

$$\begin{aligned}\mathbf{G}(y_1, y_2) &= \mathbf{C}\langle X_1, X_2\rangle (\mathbf{G}_1(y_1), \mathbf{G}_2(y_1)) \\ &= \mathbf{C}\langle X_1, X_2\rangle \left(\mathbf{F}_1\left(h_1^{-1}(y_1)\right), \mathbf{F}_2\left(h_2^{-1}(y_2)\right)\right)\end{aligned}$$

Applying an increasing transformation does not change the copula function, only the marginals

The copula function is the minimum exhaustive statistic of the dependence

Non-linear stochastic dependence

If X_1 and X_2 are two positive random variables, the previous theorem implies that:

$$\begin{aligned} \mathbf{C}\langle X_1, X_2\rangle &= \mathbf{C}\langle \ln X_1, X_2\rangle \\ &= \mathbf{C}\langle \ln X_1, \ln X_2\rangle \\ &= \mathbf{C}\langle X_1, \exp X_2\rangle \\ &= \mathbf{C}\left\langle \sqrt{X_1}, \exp X_2\right\rangle \end{aligned}$$

Concordance measures

A numeric measure m of association between X_1 and X_2 is a measure of concordance if it satisfies the following properties:

1. $-1 = m\langle X, -X\rangle \leq m\langle \mathbf{C}\rangle \leq m\langle X, X\rangle = 1$;
2. $m\left\langle \mathbf{C}^{\perp}\right\rangle = 0$;
3. $m\langle -X_1, X_2\rangle = m\langle X_1, -X_2\rangle = -m\langle X_1, X_2\rangle$;
4. if $\mathbf{C}_1 \prec \mathbf{C}_2$, then $m\langle \mathbf{C}_1\rangle \leq m\langle \mathbf{C}_2\rangle$;

We have:

$$\mathbf{C} \prec \mathbf{C}^{\perp} \Rightarrow m\langle \mathbf{C}\rangle < 0$$

and:

$$\mathbf{C} \succ \mathbf{C}^{\perp} \Rightarrow m\langle \mathbf{C}\rangle > 0$$

Kendall's tau and Spearman's rho

- Kendall's tau is the probability of concordance minus the probability of discordance:

$$\begin{aligned}\tau &= \Pr\{(X_i - X_j)\cdot(Y_i - Y_j) > 0\} - \Pr\{(X_i - X_j)\cdot(Y_i - Y_j) < 0\} \\ &= 4\iint_{[0,1]^2} \mathbf{C}(u_1, u_2)\,\mathrm{d}\mathbf{C}(u_1, u_2) - 1\end{aligned}$$

- Spearman's rho is the linear correlation of the rank statistics:

$$\begin{aligned}\varrho &= \frac{\operatorname{cov}(\mathbf{F}_X(X), \mathbf{F}_Y(Y))}{\sigma(\mathbf{F}_X(X))\cdot\sigma(\mathbf{F}_Y(Y))} \\ &= 12\iint_{[0,1]^2} u_1 u_2\,\mathrm{d}\mathbf{C}(u_1, u_2) - 3\end{aligned}$$

- For the normal copula, we have:

$$\tau = \frac{2}{\pi}\arcsin\rho \qquad \text{and} \qquad \varrho = \frac{6}{\pi}\arcsin\frac{\rho}{2}$$

Exhaustive vs non-exhaustive statistics of stochastic dependence

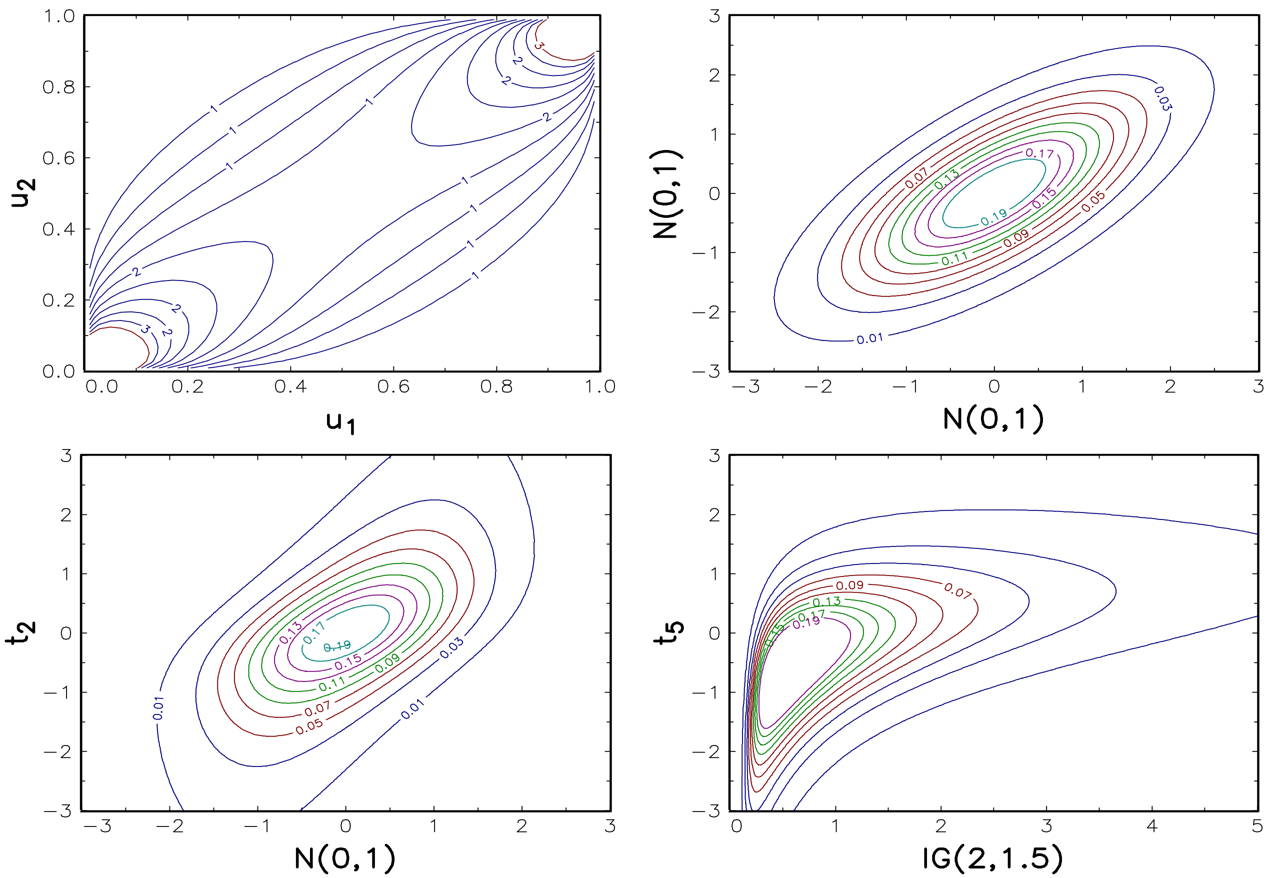


Figure: Contour lines of bivariate densities (Normal copula with $\tau = 50\%$)

Linear correlation

The linear correlation (or Pearson's correlation) is defined as follows:

$$\rho \langle X_1, X_2 \rangle = \frac{\mathbb{E}[X_1 \cdot X_2] - \mathbb{E}[X_1] \cdot \mathbb{E}[X_2]}{\sigma(X_1) \cdot \sigma(X_2)}$$

It satisfies the following properties:

- if $\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^{\perp}$, then $\rho \langle X_1, X_2 \rangle = 0$
- ρ is an increasing function with respect to the concordance measure:

$$\mathbf{C}_1 \succ \mathbf{C}_2 \Rightarrow \rho_1 \langle X_1, X_2 \rangle \geq \rho_2 \langle X_1, X_2 \rangle$$

- $\rho \langle X_1, X_2 \rangle$ is bounded:

$$\rho^- \langle X_1, X_2 \rangle \leq \rho \langle X_1, X_2 \rangle \leq \rho^+ \langle X_1, X_2 \rangle$$

  and the bounds are reached for the Fréchet copulas $\mathbf{C}^-$ and $\mathbf{C}^+$

Linear correlation

1. However, we don't have $\rho\left\langle \mathbf{C}^{-}\right\rangle = -1$ and $\rho\left\langle \mathbf{C}^{+}\right\rangle = +1$. If we use the stochastic representation of Fréchet bounds, we have:

$$\rho^{-}\left\langle X_1, X_2\right\rangle = \rho^{+}\left\langle X_1, X_2\right\rangle = \frac{\mathbb{E}\left[f_1\left(X\right)\cdot f_2\left(X\right)\right] - \mathbb{E}\left[f_1\left(X\right)\right]\cdot\mathbb{E}\left[f_2\left(X\right)\right]}{\sigma\left(f_1\left(X\right)\right)\cdot\sigma\left(f_2\left(X\right)\right)}$$

   The solution of the equation $\rho^{-}\left\langle X_1, X_2\right\rangle = -1$ is $f_1\left(x\right) = a_1 x + b_1$ and $f_2\left(x\right) = a_2 x + b_2$ where $a_1 a_2 < 0$. For the equation $\rho^{+}\left\langle X_1, X_2\right\rangle = +1$, the condition becomes $a_1 a_2 > 0$

2. Moreover, we have:

$$\rho\left\langle X_1, X_2\right\rangle = \rho\left\langle f_1\left(X_1\right), f_2\left(X_2\right)\right\rangle \Leftrightarrow \left\{\begin{array}{l} f_1\left(x\right) = a_1 x + b_1 \\ f_2\left(x\right) = a_2 x + b_2 \\ a_1 a_2 > 0 \end{array}\right.$$

Remark

The linear correlation is only valid for a linear (or Gaussian) world. **The copula function generalizes the concept of linear correlation in a non-Gaussian non-linear world**

Linear correlation

Example

We consider the bivariate log-normal random vector (X_1, X_2) where $X_1 \sim \mathcal{LN}\left(\mu_1, \sigma_1^2\right)$, $X_2 \sim \mathcal{LN}\left(\mu_2, \sigma_2^2\right)$ and $\rho = \rho \left\langle \ln X_1, \ln X_2 \right\rangle$.

We can show that:

$$\mathbb{E}\left[X_1^{p_1} \cdot X_2^{p_2}\right] = \exp\left(p_1\mu_1 + p_2\mu_2 + \frac{p_1^2\sigma_1^2 + p_2^2\sigma_2^2}{2} + p_1 p_2 \rho \sigma_1 \sigma_2\right)$$

and:

$$\rho \left\langle X_1, X_2 \right\rangle = \frac{\exp\left(\rho\sigma_1\sigma_2\right) - 1}{\sqrt{\exp\left(\sigma_1^2\right) - 1} \cdot \sqrt{\exp\left(\sigma_2^2\right) - 1}}$$

Linear correlation

If $\sigma_1 = 1$ and $\sigma_2 = 3$, we obtain the following results:

Copula	$\rho\langle X_1, X_2\rangle$	$\tau\langle X_1, X_2\rangle$	$\varrho\langle X_1, X_2\rangle$
$\mathbf{C}^-$	-0.008	-1.000	-1.000
$\rho = -0.7$	-0.007	-0.494	-0.683
$\mathbf{C}^\perp$	0.000	0.000	0.000
$\rho = 0.7$	0.061	0.494	0.683
$\mathbf{C}^+$	0.162	1.000	1.000

Tail dependence

Definition

We consider the following statistic:

$$\lambda^{+} = \lim_{u \to 1^{-}} \frac{1 - 2u + \mathbf{C}(u, u)}{1 - u}$$

We say that $\mathbf{C}$ has an upper tail dependence when $\lambda^{+} \in (0, 1]$ and $\mathbf{C}$ has no upper tail dependence when $\lambda^{+} = 0$

- For the lower tail dependence λ^{-}, the limit becomes:

$$\lambda^{-} = \lim_{u \to 0^{+}} \frac{\mathbf{C}(u, u)}{u}$$

- We notice that λ^{+} and λ^{-} can also be defined as follows:

$$\lambda^{+} = \lim_{u \to 1^{-}} \Pr\{U_2 > u \mid U_1 > u\}$$

and:

$$\lambda^{-} = \lim_{u \to 0^{+}} \Pr\{U_2 < u \mid U_1 < u\}$$

Tail dependence

- For the copula functions $\mathbf{C}^-$ and $\mathbf{C}^\perp$, we have $\lambda^- = \lambda^+ = 0$
- For the copula $\mathbf{C}^+$, we obtain $\lambda^- = \lambda^+ = 1$
- In the case of the Gumbel copula:

$$\mathbf{C}(u_1, u_2; \theta) = \exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta}\right)$$

  we obtain $\lambda^- = 0$ and $\lambda^+ = 2 - 2^{1/\theta}$
- In the case of the Clayton copula:

$$\mathbf{C}(u_1, u_2; \theta) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$$

  we obtain $\lambda^- = 2^{-1/\theta}$ and $\lambda^+ = 0$

Tail dependence

The quantile-quantile dependence function is equal to:

$$
\begin{aligned}
\lambda^{+}(\alpha) &= \Pr\left\{X_2 > \mathbf{F}_2^{-1}(\alpha) \mid X_1 > \mathbf{F}_1^{-1}(\alpha)\right\} \\
&= \frac{\Pr\left\{X_2 > \mathbf{F}_2^{-1}(\alpha), X_1 > \mathbf{F}_1^{-1}(\alpha)\right\}}{\Pr\left\{X_1 > \mathbf{F}_1^{-1}(\alpha)\right\}} \\
&= \frac{1 - \Pr\left\{X_1 \leq \mathbf{F}_1^{-1}(\alpha)\right\} - \Pr\left\{X_2 \leq \mathbf{F}_2^{-1}(\alpha)\right\}}{1 - \Pr\left\{X_1 \leq \mathbf{F}_1^{-1}(\alpha)\right\}} + \\
&\qquad \frac{\Pr\left\{X_2 \leq \mathbf{F}_2^{-1}(\alpha), X_1 \leq \mathbf{F}_1^{-1}(\alpha)\right\}}{1 - \Pr\left\{\mathbf{F}_1(X_1) \leq \alpha\right\}} \\
&= \frac{1 - 2\alpha + \mathbf{C}(\alpha, \alpha)}{1 - \alpha}
\end{aligned}
$$

Tail dependence

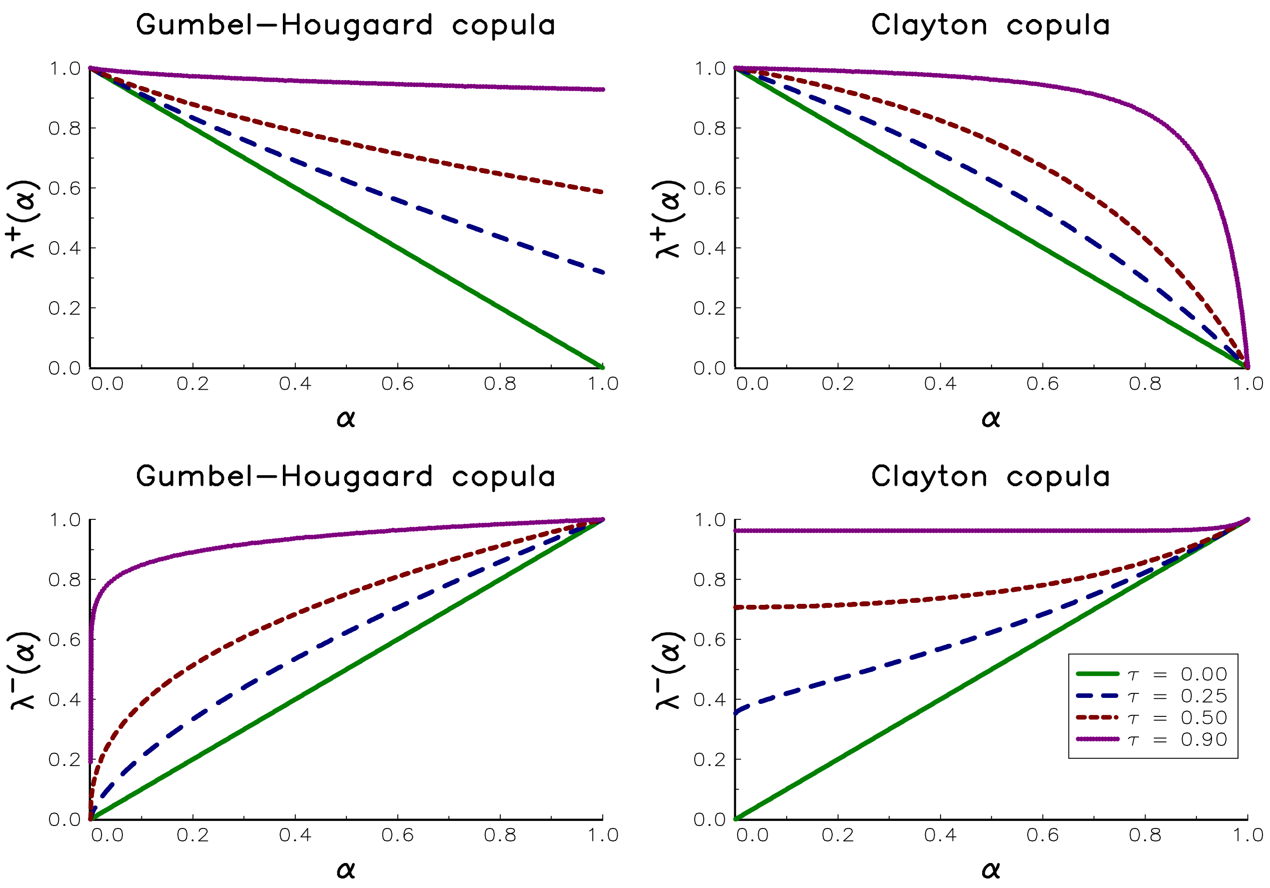


Figure: Quantile-quantile dependence measures $\lambda^{+}(\alpha)$ and $\lambda^{-}(\alpha)$

Risk interpretation of the tail dependence

We consider two portfolios, whose losses correspond to the random variables L_1 and L_2 with probability distributions $\mathbf{F}_1$ and $\mathbf{F}_2$. We have:

$$\begin{aligned}
\lambda^{+}(\alpha) &= \Pr\left\{L_2 > \mathbf{F}_2^{-1}(\alpha) \mid L_1 > \mathbf{F}_1^{-1}(\alpha)\right\} \\
&= \Pr\left\{L_2 > \mathrm{VaR}_\alpha(L_2) \mid L_1 > \mathrm{VaR}_\alpha(L_1)\right\}
\end{aligned}$$

Archimedean copulas

Definition

An Archimedean copula is defined by:

$$\mathbf{C}(u_1, u_2) = \begin{cases} \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) & \text{if } \varphi(u_1) + \varphi(u_2) \leq \varphi(0) \\ 0 & \text{otherwise} \end{cases}$$

where φ a C^2 is a function which satisfies $\varphi(1) = 0$, $\varphi'(u) < 0$ and $\varphi''(u) > 0$ for all $u \in [0, 1]$

$\Rightarrow \varphi(u)$ is called the generator of the copula function

Archimedean copulas

Example

If $\varphi(u) = u^{-1} - 1$, we have $\varphi^{-1}(u) = (1+u)^{-1}$ and:

$$\mathbf{C}(u_1, u_2) = \left(1 + \left(u_1^{-1} - 1 + u_2^{-1} - 1\right)\right)^{-1} = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

The Gumbel logistic copula is then an Archimedean copula

Remark

- *The product copula* $\mathbf{C}^{\perp}$ *is Archimedean and the associated generator is* $\varphi(u) = -\ln u$
- *Concerning Fréchet copulas, only* $\mathbf{C}^{-}$ *is Archimedean with* $\varphi(u) = 1 - u$

Archimedean copulas

Table: Archimedean copula functions

Copula	$\varphi(u)$	$\mathbf{C}(u_1, u_2)$
$\mathbf{C}^{\perp}$	$-\ln u$	$u_1 u_2$
Clayton	$u^{-\theta} - 1$	$\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$
Frank	$-\ln \dfrac{e^{-\theta u} - 1}{e^{-\theta} - 1}$	$-\dfrac{1}{\theta} \ln \left(1 + \dfrac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right)$
Gumbel	$(-\ln u)^{\theta}$	$\exp\left(-\left(\tilde{u}_1^{\theta} + \tilde{u}_2^{\theta}\right)^{1/\theta}\right)$
Joe	$-\ln\left(1 - (1 - u)^{\theta}\right)$	$1 - \left(\bar{u}_1^{\theta} + \bar{u}_2^{\theta} - \bar{u}_1^{\theta}\bar{u}_2^{\theta}\right)^{1/\theta}$

We use the notations $\bar{u} = 1 - u$ and $\tilde{u} = -\ln u$

Multivariate Normal copula

The Normal copula is the dependence function of the multivariate normal distribution with a correlation matrix ρ:

$$\mathbf{C}\left(u_1,\ldots,u_n;\rho\right)=\Phi_n\left(\Phi^{-1}\left(u_1\right),\ldots,\Phi^{-1}\left(u_n\right);\rho\right)$$

By using the canonical decomposition of the multivariate density function:

$$f\left(x_1,\ldots,x_n\right)=c\left(\mathbf{F}_1\left(x_1\right),\ldots,\mathbf{F}_n\left(x_n\right)\right)\prod_{i=1}^{n}f_i\left(x_i\right)$$

we deduce that the probability density function of the Normal copula is:

$$c\left(u_1,\ldots,u_n,;\rho\right)=\frac{1}{\left|\rho\right|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}x^\top\left(\rho^{-1}-I_n\right)x\right)$$

where $x_i=\Phi^{-1}\left(u_i\right)$

Bivariate Normal copula

In the bivariate case, we obtain:

$$c(u_1, u_2; \rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)} + \frac{x_1^2 + x_2^2}{2}\right)$$

It follows that the expression of the bivariate Normal copula function is also equal to:

$$\mathbf{C}(u_1, u_2; \rho) = \int_{-\infty}^{\mathbf{\Phi}^{-1}(u_1)} \int_{-\infty}^{\mathbf{\Phi}^{-1}(u_2)} \phi_2(x_1, x_2; \rho) \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

where $\phi_2(x_1, x_2; \rho)$ is the bivariate normal density:

$$\phi_2(x_1, x_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)}\right)$$

Bivariate Normal copula

Remark

Let (X_1, X_2) be a standardized Gaussian random vector, whose cross-correlation is ρ. Using the Cholesky decomposition, we write X_2 as follows: $X_2 = \rho X_1 + \sqrt{1-\rho^2} X_3$ where $X_3 \sim \mathcal{N}(0,1)$ is independent from X_1 and X_2. We have:

$$\begin{aligned}
\Phi_2(x_1, x_2; \rho) &= \Pr\{X_1 \leq x_1, X_2 \leq x_2\} \\
&= \mathbb{E}\left[\Pr\left\{X_1 \leq x_1, \rho X_1 + \sqrt{1-\rho^2} X_3 \leq x_2 \mid X_1\right\}\right] \\
&= \int_{-\infty}^{x_1} \Phi\left(\frac{x_2 - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x)\,\mathrm{d}x
\end{aligned}$$

It follows that:

$$\mathbf{C}(u_1, u_2; \rho) = \int_{-\infty}^{\Phi^{-1}(u_1)} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x)\,\mathrm{d}x$$

Bivariate Normal copula

- We deduce that:

$$\mathbf{C}(u_1, u_2; \rho) = \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right) \mathrm{d}u$$

- We have:

$$\tau = \frac{2}{\pi} \arcsin \rho$$

and:

$$\varrho = \frac{6}{\pi} \arcsin \frac{\rho}{2}$$

- We can show that:

$$\lambda^+ = \lambda^- = \begin{cases} 0 & \text{if } \rho < 1 \\ 1 & \text{if } \rho = 1 \end{cases}$$

Bivariate Normal copula

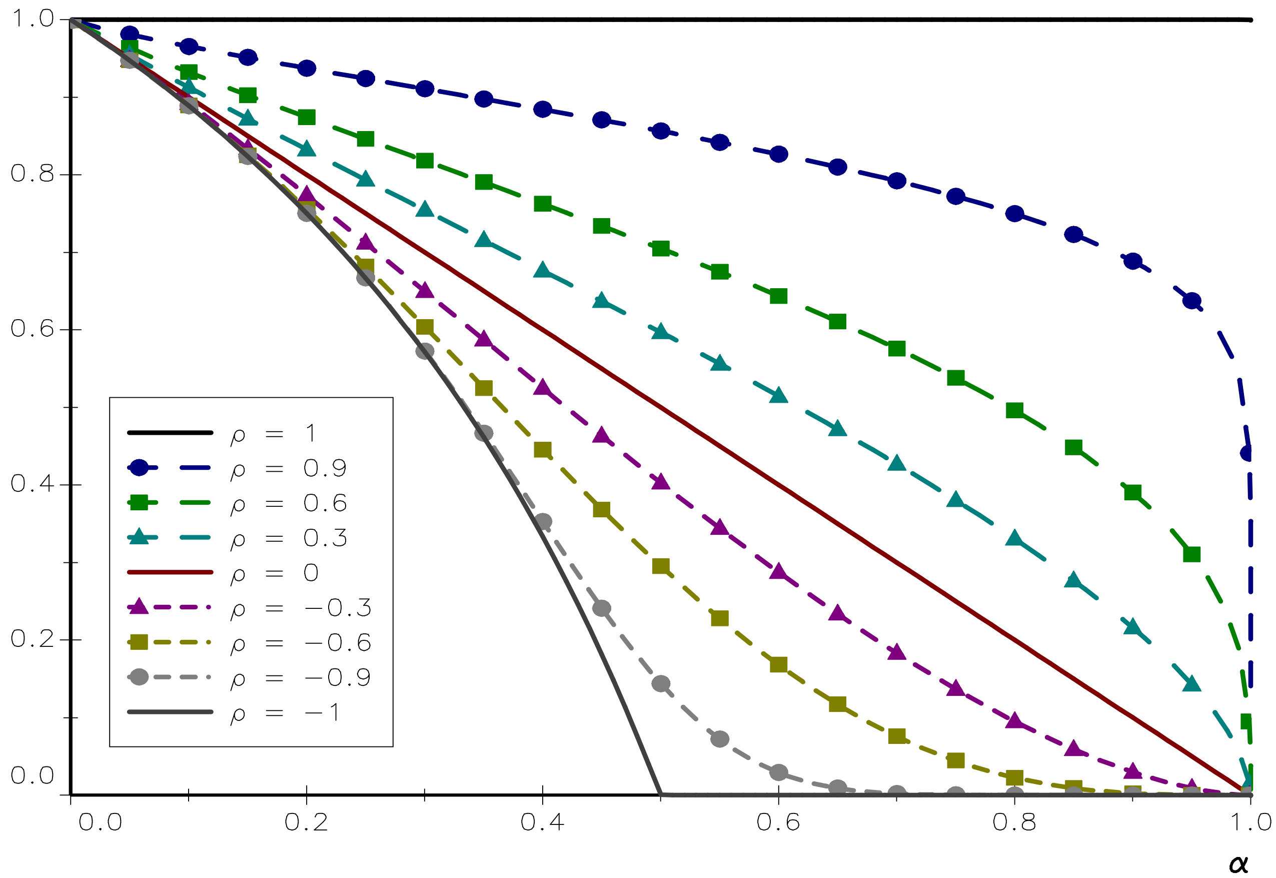


Figure: Tail dependence $\lambda^{+}(\alpha)$ for the Normal copula

Multivariate Student's t copula

We have:

$$\mathbf{C}\left(u_1,\ldots,u_n;\rho,\nu\right)=\mathbf{T}_n\left(\mathbf{T}_\nu^{-1}\left(u_1\right),\ldots,\mathbf{T}_\nu^{-1}\left(u_n\right);\rho,\nu\right)$$

By using the definition of the cumulative distribution function:

$$\mathbf{T}_n\left(x_1,\ldots,x_n;\rho,\nu\right)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\frac{\Gamma\left(\frac{\nu+n}{2}\right)\left|\rho\right|^{-\frac{1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\left(\nu\pi\right)^{\frac{n}{2}}}\left(1+\frac{1}{\nu}x^\top\rho^{-1}x\right)^{-\frac{\nu+n}{2}}\,\mathrm{d}x$$

we can show that the copula density function is then:

$$c\left(u_1,\ldots,u_n,;\rho,\nu\right)=\left|\rho\right|^{-\frac{1}{2}}\frac{\Gamma\left(\frac{\nu+n}{2}\right)\left[\Gamma\left(\frac{\nu}{2}\right)\right]^n}{\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^n\Gamma\left(\frac{\nu}{2}\right)}\frac{\left(1+\frac{1}{\nu}x^\top\rho^{-1}x\right)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^n\left(1+\frac{x_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}}$$

where $x_i=\mathbf{T}_\nu^{-1}\left(u_i\right)$

Bivariate Student's t copula

- We have:

$$\mathbf{C}\left(u_1, u_2; \rho, \nu\right) = \int_0^{u_1} \mathbf{C}_{2|1}\left(u, u_2; \rho, \nu\right) \,\mathrm{d}u$$

where:

$$\mathbf{C}_{2|1}\left(u_1, u_2; \rho, \nu\right) = \mathbf{T}_{\nu+1}\left(\left(\frac{\nu+1}{\nu+\left[\mathbf{T}_\nu^{-1}\left(u_1\right)\right]^2}\right)^{1/2} \frac{\mathbf{T}_\nu^{-1}\left(u_2\right) - \rho \mathbf{T}_\nu^{-1}\left(u_1\right)}{\sqrt{1-\rho^2}}\right)$$

- We have:

$$\lambda^+ = 2 - 2 \cdot \mathbf{T}_{\nu+1}\left(\left(\frac{\left(\nu+1\right)\left(1-\rho\right)}{\left(1+\rho\right)}\right)^{1/2}\right) = \left\{ \begin{array}{ll} 0 & \text{if } \rho = -1 \\ > 0 & \text{if } \rho > -1 \end{array} \right.$$

Bivariate Student's t copula

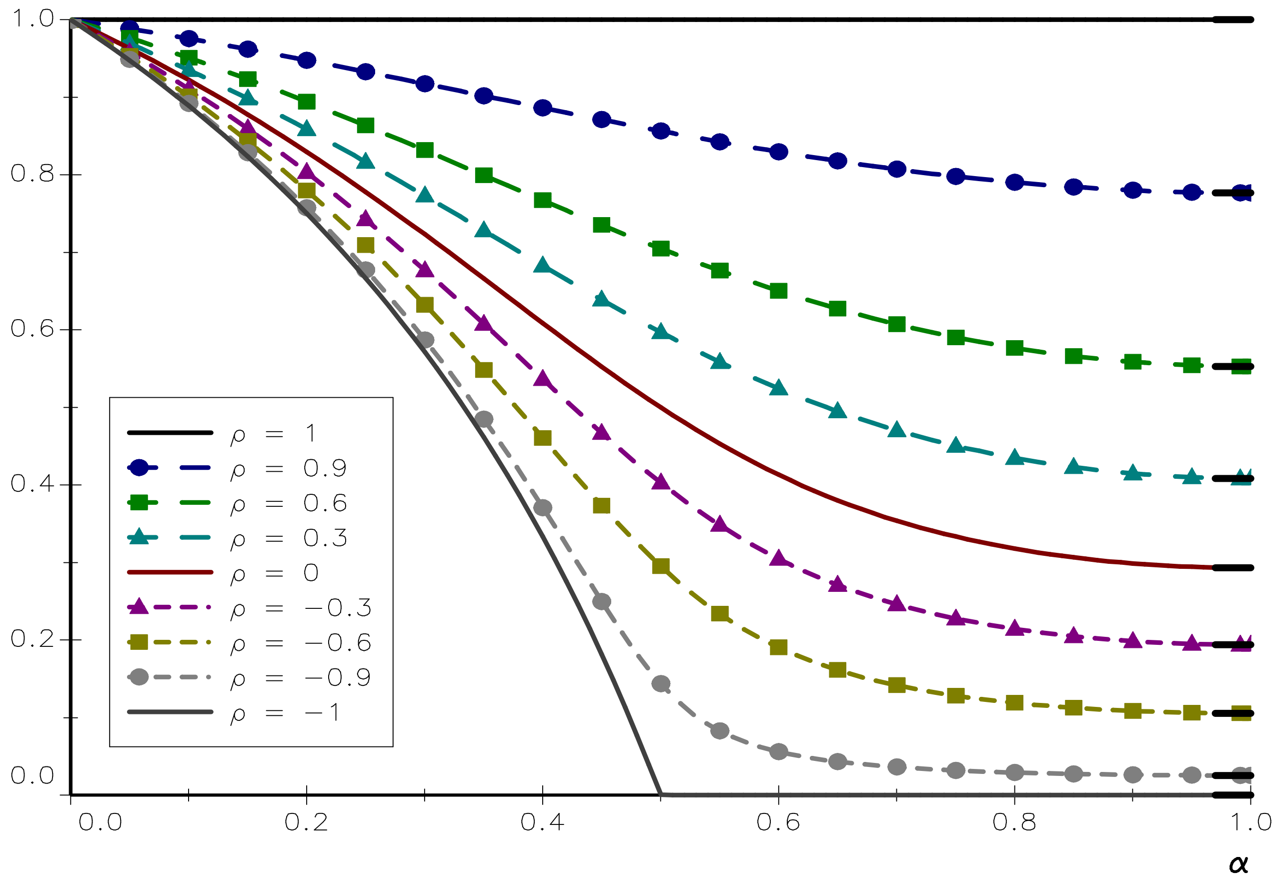


Figure: Tail dependence $\lambda^{+}(\alpha)$ for the Student's t copula $(\nu = 1)$

Bivariate Student's t copula

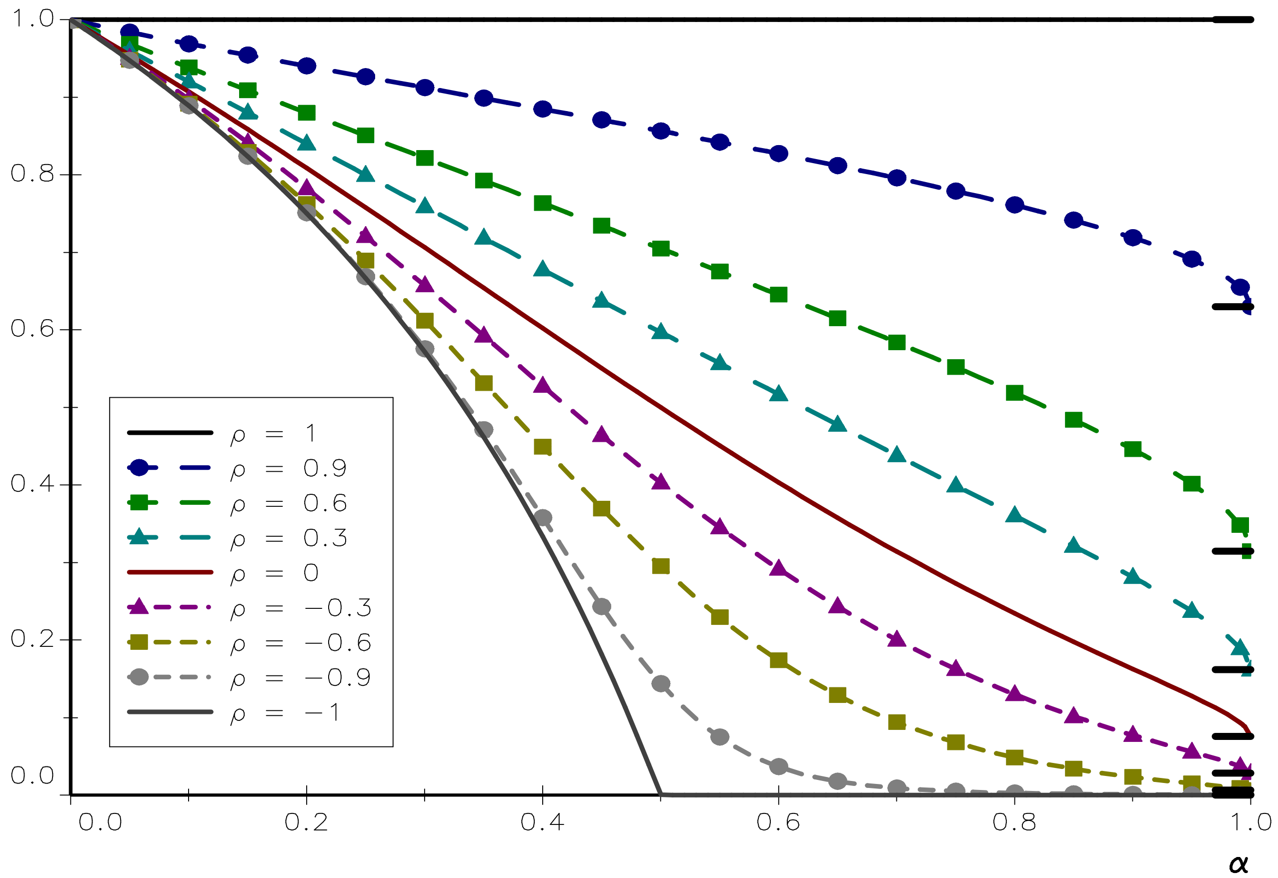


Figure: Tail dependence $\lambda^{+}(\alpha)$ for the Student's t copula ($\nu = 4$)

Dependogram

The dependogram is the scatter plot between $u_{t,1}$ and $u_{t,2}$ where:

$$u_{t,i} = \frac{1}{T+1}\mathfrak{R}_{t,i}$$

and $\mathfrak{R}_{t,i}$ is the rank statistic (T is the sample size)

Example

$x_{t,1}$	-3	4	1	8
$x_{t,2}$	105	65	17	9
$\mathfrak{R}_{t,1}$	1	3	2	4
$\mathfrak{R}_{t,2}$	4	3	2	1
$u_{t,1}$	0.20	0.60	0.40	0.80
$u_{t,2}$	0.80	0.60	0.40	0.20

Dependogram

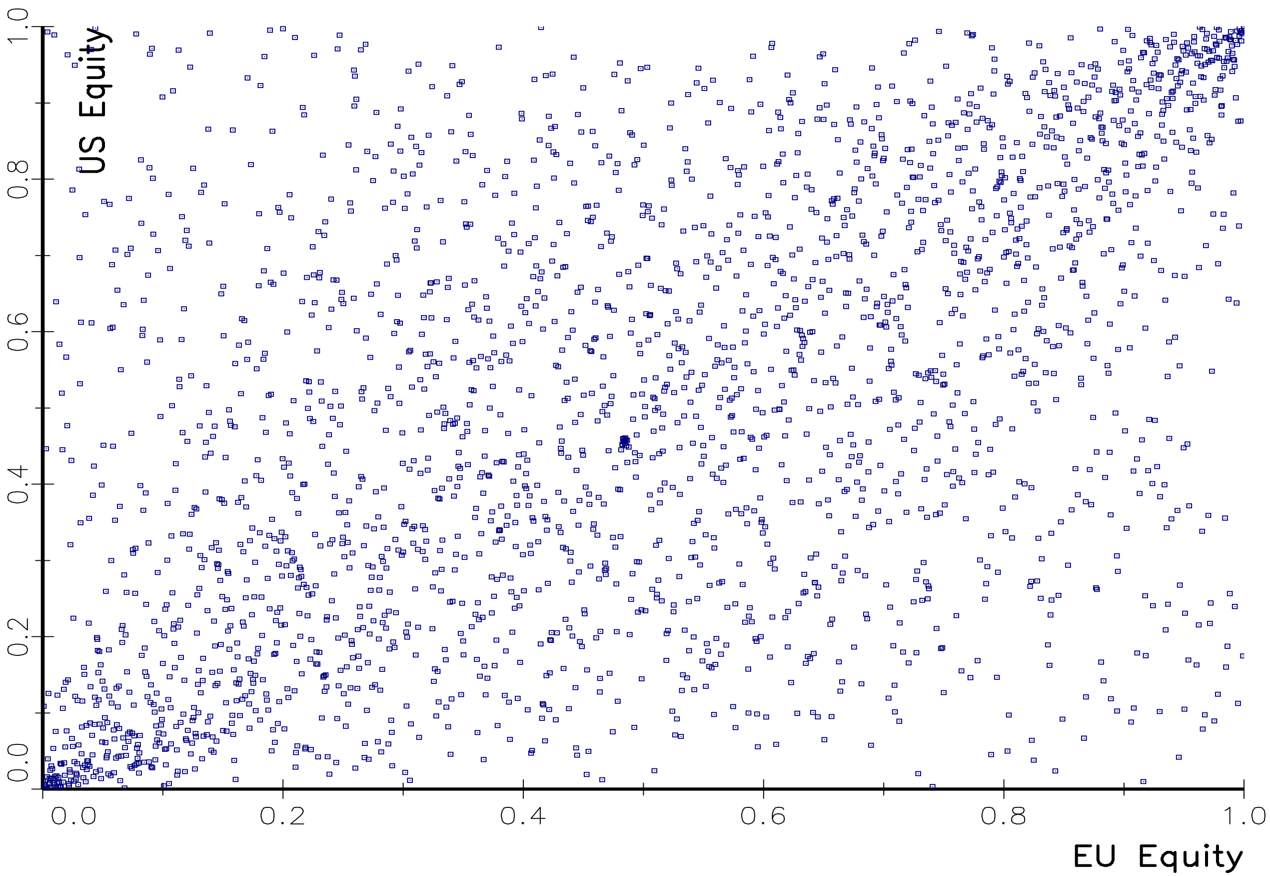


Figure: Dependogram of EU and US equity returns ($\rho = 57.8\%$)

Dependogram

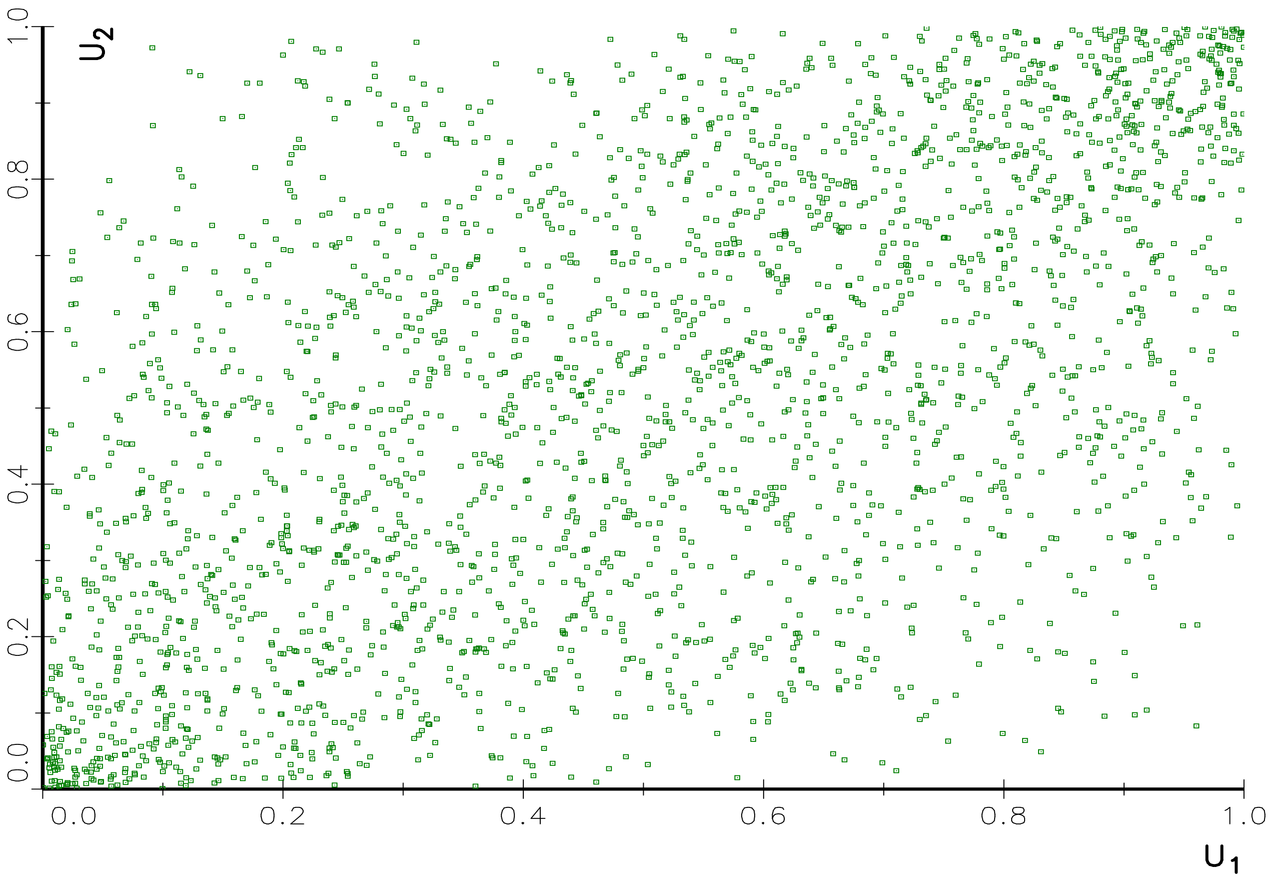


Figure: Dependogram of simulated Gaussian returns ($\rho = 57.8\%$)

The method of moments

If $\tau = f_\tau(\theta)$ is the relationship between θ and Kendall's tau, the MM estimator is simply the inverse of this relationship:

$$\hat{\theta} = f_\tau^{-1}(\hat{\tau})$$

where $\hat{\tau}$ is the estimate of Kendall's tau based on the sample

Remark

We have:

$$\hat{\tau} = \frac{c - d}{c + d}$$

where c and d are the number of concordant and discordant pairs

For instance, in the case of the Gumbel copula, we have:

$$\tau = \frac{\theta - 1}{\theta}$$

and:

$$\hat{\theta} = \frac{1}{1 - \hat{\tau}}$$

The method of maximum likelihood

We have:

$$\mathbf{F}(x_1, \ldots, x_n) = \mathbf{C}(\mathbf{F}_1(x_1; \theta_1), \ldots, \mathbf{F}_n(x_n; \theta_n); \theta_c)$$

with two types of parameters:

- the parameters $(\theta_1, \ldots, \theta_n)$ of univariate distribution functions
- the parameters θ_c of the copula function

The expression of the log-likelihood function is:

$$\begin{aligned}\ell(\theta_1, \ldots, \theta_n, \theta_c) &= \sum_{t=1}^{T} \ln c(\mathbf{F}_1(x_{t,1}; \theta_1), \ldots, \mathbf{F}_n(x_{t,n}; \theta_n); \theta_c) + \\ & \sum_{t=1}^{T} \sum_{i=1}^{n} \ln f_i(x_{t,i}; \theta_i)\end{aligned}$$

The ML estimator is then defined as follows:

$$\left(\hat{\theta}_1, \ldots, \hat{\theta}_n, \hat{\theta}_c\right) = \arg\max \ell(\theta_1, \ldots, \theta_n, \theta_c)$$

The method of inference functions for marginals

The IFM method is a two-stage parametric method:

1. the first stage involves maximum likelihood from univariate marginals
2. the second stage involves maximum likelihood of the copula parameters θ_c with the univariate parameters $\hat{\theta}_1, \ldots, \hat{\theta}_n$ held fixed from the first stage:

$$\hat{\theta}_c = \arg\max \sum_{t=1}^{T} \ln c \left(\mathbf{F}_1 \left(x_{t,1}; \hat{\theta}_1 \right), \ldots, \mathbf{F}_n \left(x_{t,n}; \hat{\theta}_n \right); \theta_c \right)$$

The omnibus method

The omnibus method replaces the marginals $\mathbf{F}_1, \ldots, \mathbf{F}_n$ by their non-parametric estimates:

$$\hat{\theta}_c = \arg\max \sum_{t=1}^{T} \ln c \left(\hat{\mathbf{F}}_1 \left(x_{t,1} \right), \ldots, \hat{\mathbf{F}}_n \left(x_{t,n} \right); \theta_c \right)$$

where:

$$\hat{\mathbf{F}}_i \left(x_{t,i} \right) = u_{t,i} = \frac{1}{T+1} \mathfrak{R}_{t,i}$$

Estimation of the Normal copula

In the case of the Normal copula, the matrix ρ of the parameters is estimated with the following algorithm:

1. we first transform the uniform variates $u_{t,i}$ into Gaussian variates:

$$n_{t,i} = \Phi^{-1}\left(u_{t,i}\right)$$

2. we then calculate the correlation matrix $\hat{\rho}$ of the Gaussian variates $n_{t,i}$.

Order statistics

Definition

- Let $X_1, \ldots, X_n$ be *iid* random variables, whose probability distribution is denoted by $\mathbf{F}$
- We rank these random variables by increasing order:

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \leq X_{n:n}$$

- $X_{i:n}$ is called the i^{th} order statistic in the sample of size n
- We note $x_{i:n}$ the corresponding random variate or the value taken by $X_{i:n}$

Order statistics

We have:

$$
\begin{aligned}
\mathbf{F}_{i:n}(x) &= \Pr\{X_{i:n} \leq x\} \\
&= \Pr\{\text{at least } i \text{ variables among } X_1, \ldots, X_n \text{ are less or equal to } x\} \\
&= \sum_{k=i}^{n} \Pr\{k \text{ variables among } X_1, \ldots, X_n \text{ are less or equal to } x\} \\
&= \sum_{k=i}^{n} \binom{n}{k} \mathbf{F}(x)^k (1 - \mathbf{F}(x))^{n-k}
\end{aligned}
$$

and:

$$
f_{i:n}(x) = \frac{\partial\, \mathbf{F}_{i:n}(x)}{\partial\, x}
$$

Order statistics

Example

If $X_1, \ldots, X_n$ follow a uniform distribution $\mathcal{U}_{[0,1]}$, we obtain:

$$\mathbf{F}_{i:n}(x) = \sum_{k=i}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \mathcal{IB}(x; i, n-i+1)$$

where $\mathcal{IB}(x; \alpha, \beta)$ is the regularized incomplete beta function:

$$\mathcal{IB}(x; \alpha, \beta) = \frac{1}{\mathfrak{B}(\alpha, \beta)} \int_0^x t^{\alpha-1} (1-t)^{\beta-1} \, \mathrm{d}t$$

We deduce that $X_{i:n} \sim \mathcal{B}(i, n-i+1)$ and[a]:

$$\mathbb{E}[X_{i:n}] = \mathbb{E}[\mathcal{B}(i, n-i+1)] = \frac{i}{n+1}$$

[a] We recall that $\mathbb{E}[\mathcal{B}(\alpha, \beta)] = \alpha / (\alpha + \beta)$

Extreme order statistics

The extreme order statistics are:

$$X_{1:n} = \min(X_1, \ldots, X_n)$$

and:

$$X_{n:n} = \max(X_1, \ldots, X_n)$$

We have:

$$\begin{aligned} \mathbf{F}_{1:n}(x) &= \sum_{k=1}^{n} \binom{n}{k} \mathbf{F}(x)^k (1 - \mathbf{F}(x))^{n-k} = 1 - \binom{n}{0} \mathbf{F}(x)^0 (1 - \mathbf{F}(x))^n \\ &= 1 - (1 - \mathbf{F}(x))^n \end{aligned}$$

and:

$$\begin{aligned} \mathbf{F}_{i:n}(x) &= \sum_{k=n}^{n} \binom{n}{k} \mathbf{F}(x)^k (1 - \mathbf{F}(x))^{n-k} = \binom{n}{n} \mathbf{F}(x)^n (1 - \mathbf{F}(x))^{n-n} \\ &= \mathbf{F}(x)^n \end{aligned}$$

Alternative proof

We have:

$$\begin{aligned}
\mathbf{F}_{1:n}(x) = \Pr\{\min(X_1,\ldots,X_n) \leq x\} &= 1 - \Pr\{\min(X_1,\ldots,X_n) \geq x\} \\
&= 1 - \Pr\{X_1 \geq x, X_2 \geq x, \ldots, X_n \geq x\} \\
&= 1 - \prod_{i=1}^{n} \Pr\{X_i \geq x\} \\
&= 1 - \prod_{i=1}^{n} \left(1 - \Pr\{X_i \leq x\}\right) \\
&= 1 - \left(1 - \mathbf{F}(x)\right)^n
\end{aligned}$$

and:

$$\begin{aligned}
\mathbf{F}_{n:n}(x) = \Pr\{\max(X_1,\ldots,X_n) \leq x\} &= \Pr\{X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x\} \\
&= \prod_{i=1}^{n} \Pr\{X_i \leq x\} \\
&= \mathbf{F}(x)^n
\end{aligned}$$

Extreme order statistics

We deduce that the density functions are equal to:

$$f_{1:n}(x) = n(1 - \mathbf{F}(x))^{n-1} f(x)$$

and

$$f_{n:n}(x) = n\mathbf{F}(x)^{n-1} f(x)$$

Extreme order statistics

We consider the daily returns of the MSCI USA index from 1995 to 2015

$\mathcal{H}_1$ Daily returns are Gaussian, meaning that:

$$R_t = \hat{\mu} + \hat{\sigma} X_t$$

where $X_t \sim \mathcal{N}(0,1)$, $\hat{\mu}$ is the empirical mean of daily returns and $\hat{\sigma}$ is the daily standard deviation

$\mathcal{H}_2$ Daily returns follow a Student's t distribution[19]:

$$R_t = \hat{\mu} + \hat{\sigma} \sqrt{\frac{\nu - 2}{\nu}} X_t$$

where $X_t \sim \mathbf{t}_\nu$. We consider two alternative assumptions: $\mathcal{H}_{2a} : \nu = 3$ and $\mathcal{H}_{2b} : \nu = 6$

[19] We add the factor $\sqrt{(\nu - 2)/\nu}$ in order to verify that $\operatorname{var}(R_t) = \hat{\sigma}^2$

Extreme order statistics

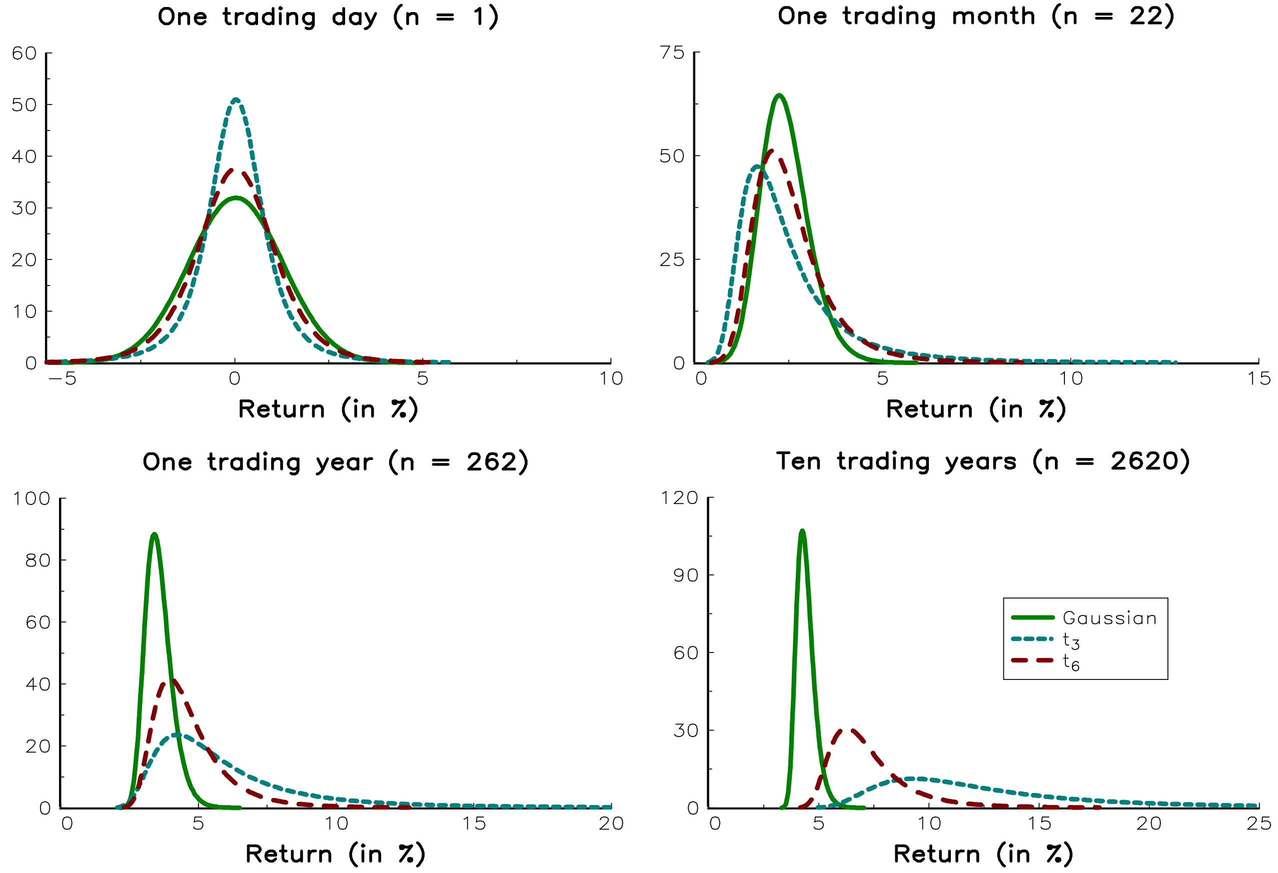


Figure: Density function of the maximum order statistic (daily return of the MSCI USA index, 1995-2015)

Extreme order statistics

Remark

The limit distributions of minima and maxima are degenerate:

$$\lim_{n\to\infty} \mathbf{F}_{1:n}(x) = \lim_{n\to\infty} 1 - (1 - \mathbf{F}(x))^n = \begin{cases} 0 & \text{if } \mathbf{F}(x) = 0 \\ 1 & \text{if } \mathbf{F}(x) > 0 \end{cases}$$

and:

$$\lim_{n\to\infty} \mathbf{F}_{n:n}(x) = \lim_{n\to\infty} \mathbf{F}(x)^n = \begin{cases} 0 & \text{if } \mathbf{F}(x) < 1 \\ 1 & \text{if } \mathbf{F}(x) = 1 \end{cases}$$

Remark

We only consider the largest order statistic $X_{n:n}$ because the minimum order statistic $X_{1:n}$ is equal to $Y_{n:n}$ by setting $Y_i = -X_i$

Univariate extreme value theory

Fisher-Tippet theorem

Let $X_1, \ldots, X_n$ be a sequence of *iid* random variables, whose distribution function is $\mathbf{F}$. If there exist two constants a_n and b_n and a non-degenerate distribution function $\mathbf{G}$ such that:

$$\lim_{n\to\infty} \Pr\left\{\frac{X_{n:n} - b_n}{a_n} \leq x\right\} = \mathbf{G}(x)$$

then $\mathbf{G}$ can be classified as one of the following three types:

Type I	(Gumbel)	$\boldsymbol{\Lambda}(x) = \exp\left(-e^{-x}\right)$
Type II	(Fréchet)	$\boldsymbol{\Phi}_\alpha(x) = \mathbb{1}(x \geq 0) \cdot \exp\left(-x^{-\alpha}\right)$
Type III	(Weibull	$\boldsymbol{\Psi}_\alpha(x) = \mathbb{1}(x \leq 0) \cdot \exp\left(-(-x)^{\alpha}\right)$

$\boldsymbol{\Lambda}$, $\boldsymbol{\Phi}_\alpha$ and $\boldsymbol{\Psi}_\alpha$ are called extreme value distributions

Fisher-Tippet theorem $\approx$ an extreme value analog of the central limit theorem

Univariate extreme value theory

We recall that:

$$\lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \exp(x)$$

Univariate extreme value theory

- We consider the exponential distribution: $\mathbf{F}(x) = 1 - \exp(-\lambda x)$. We have:

$$\lim_{n\to\infty} \mathbf{F}_{n:n}(x) = \lim_{n\to\infty} \left(1 - e^{-\lambda x}\right)^n = \lim_{n\to\infty} \left(1 - \frac{ne^{-\lambda x}}{n}\right)^n$$
$$= \lim_{n\to\infty} \exp\left(-ne^{-\lambda x}\right) = 0$$

We verify that the limit distribution is degenerate

- If we consider the affine transformation with $a_n = 1/\lambda$ et $b_n = (\ln n)/\lambda$, we obtain:

$$\Pr\left\{\frac{X_{n:n} - b_n}{a_n} \leq x\right\} = \Pr\left\{X_{n:n} \leq a_n x + b_n\right\} = \left(1 - e^{-\lambda(a_n x + b_n)}\right)^n$$
$$= \left(1 - e^{-x - \ln n}\right)^n = \left(1 - \frac{e^{-x}}{n}\right)^n$$

and:

$$\mathbf{G}(x) = \lim_{n\to\infty} \left(1 - \frac{e^{-x}}{n}\right)^n = \exp\left(-e^{-x}\right) = \mathbf{\Lambda}(x)$$

Generalized extreme value distribution

- We combine the three distributions $\mathbf{\Lambda}$, $\mathbf{\Phi}_\alpha$ et $\mathbf{\Psi}_\alpha$ into a single distribution function $\mathcal{GEV}(\mu, \sigma, \xi)$:

$$\mathbf{G}(x) = \exp\left(-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

  defined on the support $\Delta = \left\{x : 1 + \xi\sigma^{-1}(x - \mu) > 0\right\}$
- the limit case $\xi \to 0$ corresponds to the Gumbel distribution $\mathbf{\Lambda}$
- $\xi = -\alpha^{-1} > 0$ defines the Fréchet distribution $\mathbf{\Phi}_\alpha$
- the Weibull distribution $\mathbf{\Psi}_\alpha$ is obtained by considering $\xi = -\alpha^{-1} < 0$

Generalized extreme value distribution

The density function is equal to:

$$g(x) = \frac{1}{\sigma}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-(1+\xi)/\xi} \exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

Block maxima approach

The log-likelihood function is equal to:

$$\ell_t = -\ln\sigma - \left(\frac{1+\xi}{\xi}\right)\ln\left(1+\xi\left(\frac{x_t-\mu}{\sigma}\right)\right) - \left(1+\xi\left(\frac{x_t-\mu}{\sigma}\right)\right)^{-1/\xi}$$

where x_t is the observed maximum for the t^{th} period (or block maximum)

Generalized extreme value distribution

- We consider the example of the MSCI USA index
- Using daily returns, we calculate the block maximum for each period of 22 trading days and estimate the GEV distribution using the method of maximum likelihood
- We compare the estimated GEV distribution with the distribution function $\mathbf{F}_{22:22}(x)$ when we assume that daily returns are Gaussian:

α	90%	95%	96%	97%	98%	99%
Gaussian	3.26%	3.56%	3.65%	3.76%	3.92%	4.17%
GEV	3.66%	4.84%	5.28%	5.91%	6.92%	9.03%

Generalized extreme value distribution

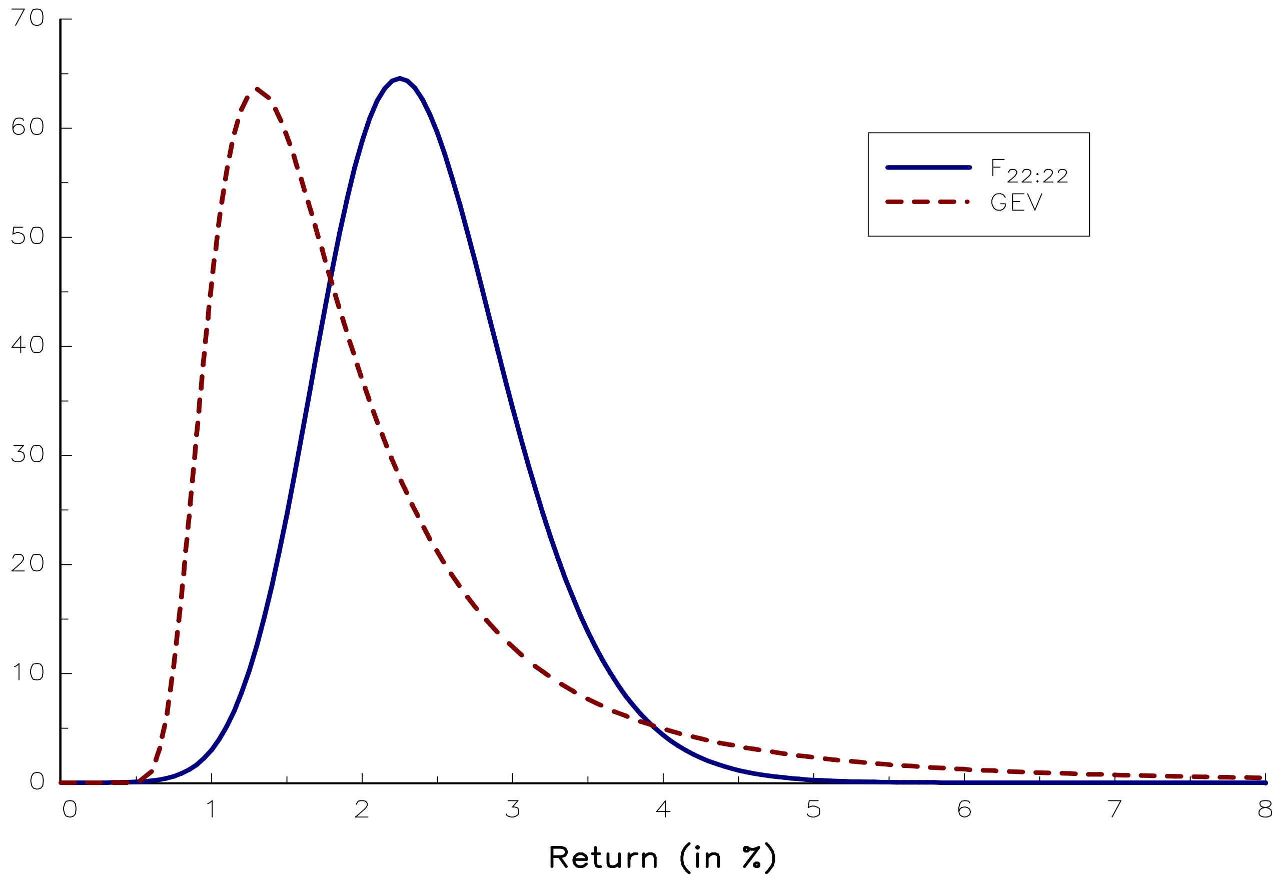


Figure: Probability density function of the maximum return $R_{22:22}$

Value-at-risk estimation

We recall that the P&L between t and $t+1$ is equal to:

$$\Pi(w) = P_{t+1}(w) - P_t(w) = P_t(w) \cdot R(w)$$

We have:

$$\mathrm{VaR}_\alpha(w) = -P_t(w) \cdot \hat{\mathbf{F}}^{-1}(1-\alpha)$$

We now estimate the GEV distribution $\hat{\mathbf{G}}$ of the maximum of $-R(w)$ for a period of n trading days. The confidence level must be adjusted in order to obtain the same return time:

$$\frac{1}{1-\alpha} \times 1 \text{ day} = \frac{1}{1-\alpha_{\mathrm{GEV}}} \times n \text{ days} \Leftrightarrow \alpha_{\mathrm{GEV}} = 1 - (1-\alpha) \cdot n$$

It follows that the value-at-risk is equal to:

$$\mathrm{VaR}_\alpha(w) = P(t) \cdot \hat{\mathbf{G}}^{-1}(\alpha_{\mathrm{GEV}}) = P(t) \cdot \left(\hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left(1 - (-\ln \alpha_{\mathrm{GEV}})^{-\hat{\xi}}\right)\right)$$

because we have $\mathbf{G}^{-1}(\alpha) = \mu - \frac{\sigma}{\xi}\left(1 - (-\ln \alpha)^{-\xi}\right)$

Value-at-risk estimation

Table: Comparing Gaussian, historical and GEV value-at-risk measures

VaR	α	Long US	Long EM	Long US Short EM	Long EM Short US
Gaussian	99.0%	2.88%	2.83%	3.06%	3.03%
	99.5%	3.19%	3.14%	3.39%	3.36%
	99.9%	3.83%	3.77%	4.06%	4.03%
Historical	99.0%	3.46%	3.61%	3.37%	3.81%
	99.5%	4.66%	4.73%	3.99%	4.74%
	99.9%	7.74%	7.87%	6.45%	7.27%
GEV	99.0%	2.64%	2.61%	2.72%	2.93%
	99.5%	3.48%	3.46%	3.41%	3.82%
	99.9%	5.91%	6.05%	5.35%	6.60%

Expected shortfall estimation

We use the peak over threshold approach (HFRM, pages 773-777)

Extreme value copulas

Definition

An extreme value (EV) copula satisfies the following relationship:

$$\mathbf{C}\left(u_1^t, \ldots, u_n^t\right) = \mathbf{C}^t\left(u_1, \ldots, u_n\right)$$

for all $t > 0$

Extreme value copulas

The Gumbel copula is an EV copula:

$$
\begin{aligned}
\mathbf{C}\left(u_1^t, u_2^t\right) &= \exp\left(-\left(\left(-\ln u_1^t\right)^\theta + \left(-\ln u_2^t\right)^\theta\right)^{1/\theta}\right) \\
&= \exp\left(-\left(t^\theta\left(\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right)\right)^{1/\theta}\right) \\
&= \left(\exp\left(-\left(\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right)^{1/\theta}\right)\right)^t \\
&= \mathbf{C}^t\left(u_1, u_2\right)
\end{aligned}
$$

Extreme value copulas

The Farlie-Gumbel-Morgenstern copula is not an EV copula:

$$\begin{aligned}
\mathbf{C}\left(u_1^t, u_2^t\right) &= u_1^t u_2^t + \theta u_1^t u_2^t \left(1 - u_1^t\right)\left(1 - u_2^t\right) \\
&= u_1^t u_2^t \left(1 + \theta - \theta u_1^t - \theta u_2^t + \theta u_1^t u_2^t\right) \\
&\neq u_1^t u_2^t \left(1 + \theta - \theta u_1 - \theta u_2 + \theta u_1 u_2\right)^t \\
&\neq \mathbf{C}^t\left(u_1, u_2\right)
\end{aligned}$$

Extreme value copulas

Show that:

- $\mathbf{C}^{+}$ is an EV copula
- $\mathbf{C}^{\perp}$ is an EV copula
- $\mathbf{C}^{-}$ is not an EV copula

Multivariate extreme value theory

Let $X = (X_1, \ldots, X_n)$ be a random vector of dimension n. We note $X_{m:m}$ the random vector of maxima:

$$X_{m:m} = \begin{pmatrix} X_{m:m,1} \\ \vdots \\ X_{m:m,n} \end{pmatrix}$$

and $\mathbf{F}_{m:m}$ the corresponding distribution function:

$$\mathbf{F}_{m:m}(x_1, \ldots, x_n) = \Pr\{X_{m:m,1} \leq x_1, \ldots, X_{m:m,n} \leq x_n\}$$

The multivariate extreme value (MEV) theory considers the asymptotic behavior of the non-degenerate distribution function $\mathbf{G}$ such that:

$$\lim_{m\to\infty} \Pr\left(\frac{X_{m:m,1} - b_{m,1}}{a_{m,1}} \leq x_1, \ldots, \frac{X_{m:m,n} - b_{m,n}}{a_{m,n}} \leq x_n\right) = \mathbf{G}(x_1, \ldots, x_n)$$

Multivariate extreme value theory

Using Sklar's theorem, there exists a copula function $\mathbf{C}\langle\mathbf{G}\rangle$ such that:

$$\mathbf{G}(x_1,\ldots,x_n) = \mathbf{C}\langle\mathbf{G}\rangle\left(\mathbf{G}_1(x_1),\ldots,\mathbf{G}_n(x_n)\right)$$

We have:

- The marginals $\mathbf{G}_1,\ldots,\mathbf{G}_n$ satisfy the Fisher-Tippet theorem
- $\mathbf{C}\langle\mathbf{G}\rangle$ is an extreme value copula

Remark

An extreme value copula satisfies the PQD property:

$$\mathbf{C}^{\perp} \prec \mathbf{C} \prec \mathbf{C}^{+}$$

Tail dependence of extreme values

We can show that the (upper) tail dependence of $\mathbf{C}\langle\mathbf{G}\rangle$ is equal to the (upper) tail dependence of $\mathbf{C}\langle\mathbf{F}\rangle$:

$$\lambda^{+}\left(\mathbf{C}\langle\mathbf{G}\rangle\right)=\lambda^{+}\left(\mathbf{C}\langle\mathbf{F}\rangle\right)$$

$\Rightarrow$ Extreme values are independent if the copula function $\mathbf{C}\langle\mathbf{F}\rangle$ has no (upper) tail dependence

Advanced topics

- Maximum domain of attraction
  - Univariate extreme value theory (HFRM, pages 765-770)
  - Multivariate extreme value theory (HFRM, pages 779 and 781-782)
- Deheuvels-Pickands representation (HFRM, pages 779-781)
- Generalized Pareto distribution $\mathcal{GPD}(\sigma, \xi)$ (HFRM, pages 773-777)

Exercises

- Copulas
  - Exercise 11.5.5 – Correlated loss given default rates
  - Exercise 11.5.6 – Calculation of correlation bounds
  - Exercise 11.5.7 – The bivariate Pareto copula
- Extreme value theory
  - Exercise 12.4.2 – Order statistics and return period
  - Exercise 12.4.4 – Extreme value theory in the bivariate case
  - Exercise 12.4.5 – Maximum domain of attraction in the bivariate case

References

Nelsen, R.B. (2006)
An Introduction to Copulas, Second edition, Springer.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapters 11 and 12.

Roncalli, T. (2020)
Handbook of Financial Risk Management – Companion Book, Chapters 11 and 12.

Course 2023-2024 in Financial Risk Management
Lecture 10. Monte Carlo Simulation Methods

Thierry Roncalli*

*Amundi Asset Management[20]

*University of Paris-Saclay

September 2023

[20]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

Uniform random numbers

The idea is to build a pseudorandom sequence $\mathcal{S}$ and repeat this sequence as often as necessary

Linear congruential generator

- The most famous and used algorithm is the linear congruential generator (LCG):

$$\begin{aligned} x_n &= (a \cdot x_{n-1} + c) \bmod m \\ u_n &= x_n / m \end{aligned}$$

  where:

  - a is the multiplicative constant
  - c is the additive constant
  - m is the modulus (or the order of the congruence)

- The initial number x_0 is called the seed
- $\{x_1, x_2, \ldots, x_n\}$ is a sequence of pseudorandom integer numbers ($0 \leq x_n < m$)
- $\{u_1, u_2, \ldots, u_n\}$ is a sequence of uniform random variates
- The maximum period is m

Linear congruential generator

Example #1

If we consider that $a = 3$, $c = 0$, $m = 11$ and $x_0 = 1$, we obtain the following sequence:

$$\{1, 3, 9, 5, 4, 1, 3, 9, 5, 4, 1, 3, 9, 5, 4, \ldots\}$$

The period length is only five, meaning that only five uniform random variates can be generated: 0.09091, 0.27273, 0.81818, 0.45455 and 0.36364

Linear congruential generator

The minimal standard LCG proposed by Lewis *et al.* (1969) is defined by $a = 7^5$, $c = 0$ and $m = 2^{31} - 1$
Its period length is equal to $m - 1 = 2^{31} - 2 \approx 2.15 \times 10^9$

Table: Simulation of 10 uniform pseudorandom numbers

n	x_n	u_n	x_n	u_n
0	1	0.000000	123 456	0.000057
1	16 807	0.000008	2 074 924 992	0.966212
2	282 475 249	0.131538	277 396 911	0.129173
3	1 622 650 073	0.755605	22 885 540	0.010657
4	984 943 658	0.458650	237 697 967	0.110687
5	1 144 108 930	0.532767	670 147 949	0.312062
6	470 211 272	0.218959	1 772 333 975	0.825307
7	101 027 544	0.047045	2 018 933 935	0.940139
8	1 457 850 878	0.678865	1 981 022 945	0.922486
9	1 458 777 923	0.679296	466 173 527	0.217079
10	2 007 237 709	0.934693	958 124 033	0.446161

Linear congruential generator

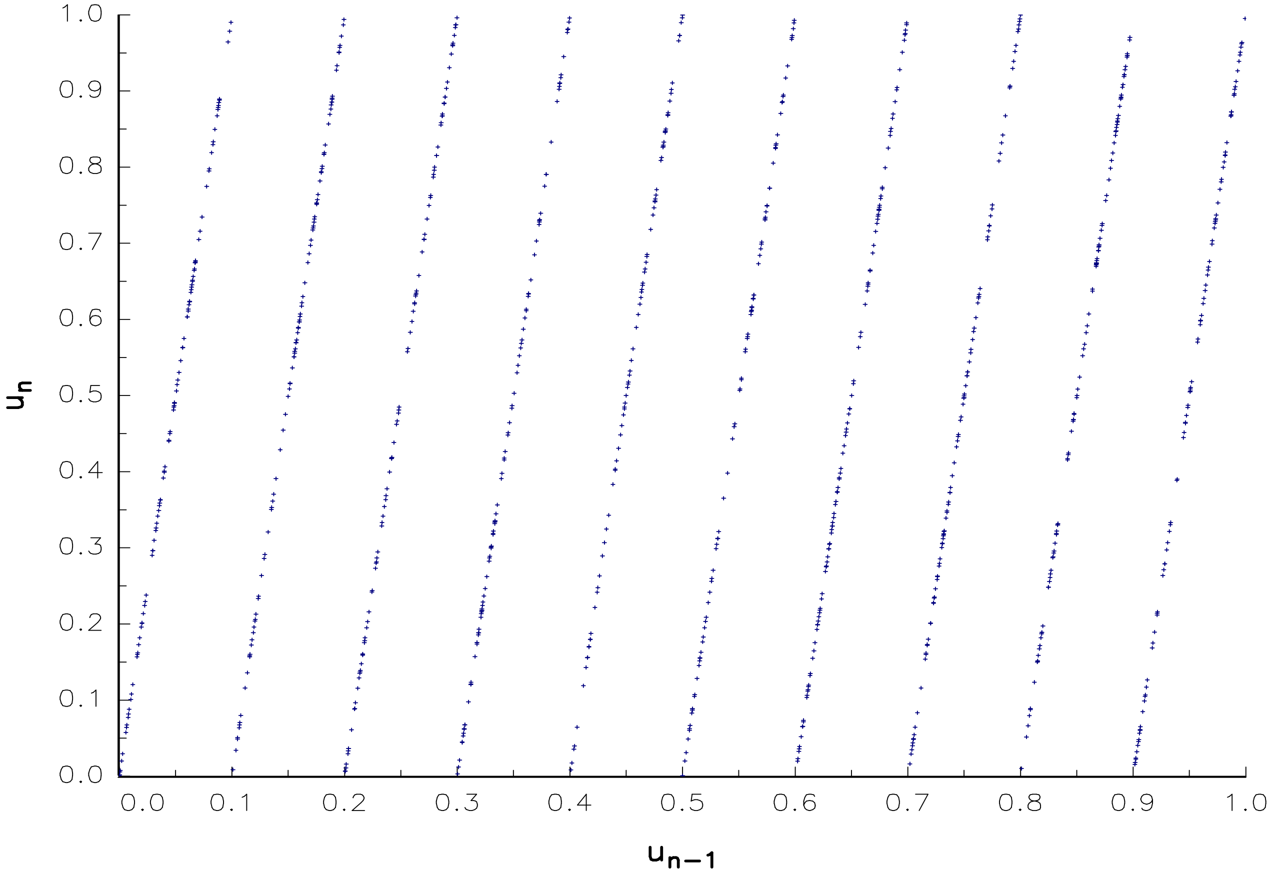


Figure: Lattice structure of the linear congruential generator

Multiple recursive generator

- We have

$$x_n = \left(\sum_{i=1}^{k} a_i \cdot x_{n-i} + c \right) \mod m$$

- The famous MRG32k3a generator of L'Ecuyer (1999) uses two 32-bit multiple recursive generators:

$$\begin{cases} x_n = (1403580 \cdot x_{n-2} - 810728 \cdot x_{n-3}) \mod m_1 \\ y_n = (527612 \cdot y_{n-1} - 1370589 \cdot y_{n-3}) \mod m_2 \end{cases}$$

  where $m_1 = 2^{32} - 209$ and $m_2 = 2^{32} - 22853$. The uniform random variate is then equal to:

$$u_n = \frac{x_n - y_n + \mathbb{1}\{x_n \leq y_n\} \cdot m_1}{m_1 + 1}$$

- The period length of this generator is equal to $2^{191} \approx 3 \times 10^{57}$

We now consider X a random variable whose distribution function is noted $\mathbf{F}$. There are many ways to simulate X, but all of them are based on uniform random variates

Method of inversion
Continuous random variables

- We assume that $\mathbf{F}$ is continuous
- Let $Y = \mathbf{F}(X)$ be the integral transform of X
- Its cumulative distribution function $\mathbf{G}$ is equal to:

$$\begin{aligned}\mathbf{G}(y) &= \Pr\{Y \leq y\} \\ &= \Pr\{\mathbf{F}(X) \leq y\} \\ &= \Pr\{X \leq \mathbf{F}^{-1}(y)\} \\ &= \mathbf{F}\left(\mathbf{F}^{-1}(y)\right) \\ &= y\end{aligned}$$

where $\mathbf{G}(0) = 0$ and $\mathbf{G}(1) = 1$

Method of inversion

Continuous random variables

- We deduce that $\mathbf{F}(X)$ has a uniform distribution $\mathcal{U}_{[0,1]}$:

$$\mathbf{F}(X) \sim \mathcal{U}_{[0,1]}$$

  IIf U is a uniform random variable, then $\mathbf{F}^{-1}(U)$ is a random variable whose distribution function is $\mathbf{F}$:

$$U \sim \mathcal{U}_{[0,1]} \Rightarrow \mathbf{F}^{-1}(U) \sim \mathbf{F}$$

- To simulate a sequence of random variates $\{x_1, \ldots, x_n\}$, we can simulate a sequence of uniform random variates $\{u_1, \ldots, u_n\}$ and apply the transform $x_i \leftarrow \mathbf{F}^{-1}(u_i)$

Method of inversion

Continuous random variables

Example #2

If we consider the generalized uniform distribution $\mathcal{U}_{[a,b]}$, we have $\mathbf{F}(x) = (x - a) / (b - a)$ and $\mathbf{F}^{-1}(u) = a + (b - a)\, u$. The simulation of random variates x_i is deduced from the uniform random variates u_i by using the following transform:

$$x_i \leftarrow a + (b - a)\, u_i$$

Method of inversion

Continuous random variables

Example #3

In the case of the exponential distribution $\mathcal{E}(\lambda)$, we have $\mathbf{F}(x) = 1 - \exp(-\lambda x)$. We deduce that:

$$x_i \leftarrow -\frac{\ln(1 - u_i)}{\lambda}$$

Since $1 - U$ is also a uniform distributed random variable, we have:

$$x_i \leftarrow -\frac{\ln(u_i)}{\lambda}$$

Method of inversion

Continuous random variables

Example #4

In the case of the Pareto distribution $\mathcal{P}(\alpha, x_-)$, we have $\mathbf{F}(x) = 1 - (x/x_-)^{-\alpha}$ and $\mathbf{F}^{-1}(u) = x_-(1-u)^{-1/\alpha}$. We deduce that:

$$x_i \leftarrow \frac{x_-}{(1-u_i)^{1/\alpha}}$$

Method of inversion

Continuous random variables

- The method of inversion is easy to implement when we know the analytical expression of $\mathbf{F}^{-1}$
- When it is not the case, we use the Newton-Raphson algorithm:

$$x_i^{m+1} = x_i^m + \frac{u_i - \mathbf{F}\left(x_i^m\right)}{f\left(x_i^m\right)}$$

  where x_i^m is the solution of the equation $\mathbf{F}(x) = u$ at the iteration m
- If we apply this algorithm to the Gaussian distribution $\mathcal{N}(0,1)$, we have:

$$x_i^{m+1} = x_i^m + \frac{u_i - \Phi\left(x_i^m\right)}{\phi\left(x_i^m\right)}$$

Method of inversion
Discrete random variables

In the case of a discrete probability distribution $\{(x_1, p_1), (x_2, p_2), \ldots, (x_n, p_n)\}$ where $x_1 < x_2 < \ldots < x_n$, we have:

$$\mathbf{F}^{-1}(u) = \begin{cases} x_1 & \text{if} \quad 0 \leq u \leq p_1 \\ x_2 & \text{if} \quad p_1 < u \leq p_1 + p_2 \\ & \vdots \\ x_n & \text{if} \quad \sum_{k=1}^{n-1} p_k < u \leq 1 \end{cases}$$

Method of inversion
Discrete random variables

- We assume that:

x_i	1	2	4	6	7	9	10
p_i	10%	20%	10%	5%	20%	30%	5%
$\mathbf{F}(x_i)$	10%	30%	40%	45%	65%	95%	100%

- The inverse function is a step function
- If $u = 0.5517$, Then $X = \mathbf{F}^{-1}(u) = \mathbf{F}^{-1}(0.5517) = 7$

Method of inversion

Discrete random variables

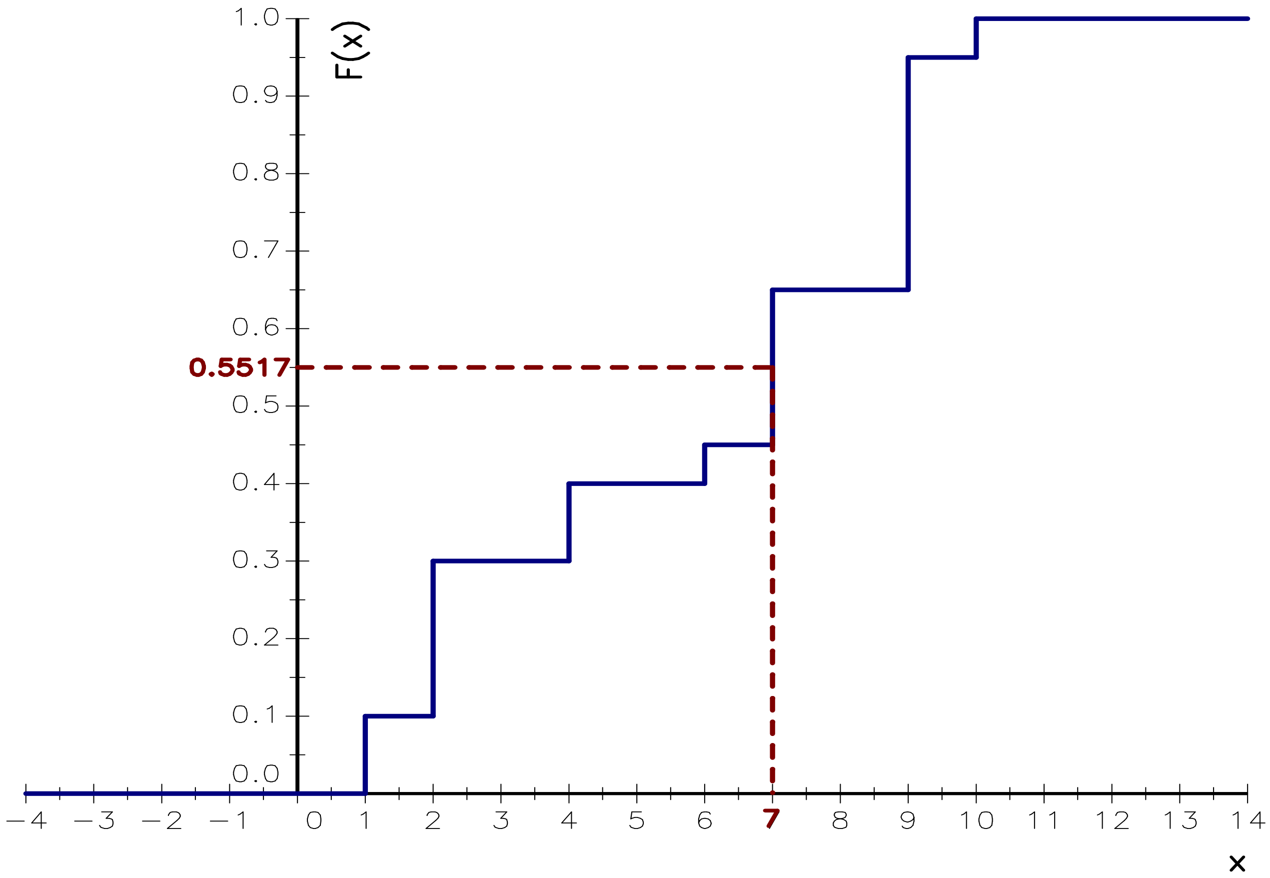


Figure: Inversion method when X is a discrete random variable

Method of inversion

Discrete random variables

Example #5

If we apply the method of inversion to the Bernoulli distribution $\mathcal{B}(p)$, we have:

$$x \leftarrow \begin{cases} 0 & \text{if} \quad 0 \leq u \leq 1-p \\ 1 & \text{if} \quad 1-p < u \leq 1 \end{cases}$$

or:

$$x \leftarrow \begin{cases} 1 & \text{if} \quad u \leq p \\ 0 & \text{if} \quad u > p \end{cases}$$

Method of inversion
Piecewise distribution functions

- A piecewise distribution function is defined as follows:

$$\mathbf{F}(x) = \mathbf{F}_m(x) \qquad \text{if } x \in \left]x^\star_{m-1}, x^\star_m\right]$$

where $x^\star_m$ are the knots of the piecewise function and:

$$\mathbf{F}_{m+1}(x^\star_m) = \mathbf{F}_m(x^\star_m)$$

- In this case, the simulated value x_i is obtained using a search algorithm:

$$x_i \leftarrow \mathbf{F}^{-1}_m(u_i) \qquad \text{if } \mathbf{F}\left(x^\star_{m-1}\right) < u_i \leq \mathbf{F}(x^\star_m)$$

Method of inversion

Piecewise distribution functions

- We consider the piecewise exponential model
- The survival function has the following expression:

$$\mathbf{S}(t) = \mathbf{S}\left(t_{m-1}^{\star}\right) e^{-\lambda_m \left(t - t_{m-1}^{\star}\right)} \qquad \text{if } t \in \left]t_{m-1}^{\star}, t_m^{\star}\right]$$

- We know that $\mathbf{S}(\boldsymbol{\tau}) \sim U$
- It follows that:

$$t_i \leftarrow t_{m-1}^{\star} + \frac{1}{\lambda_m} \ln \frac{\mathbf{S}\left(t_{m-1}^{\star}\right)}{u_i} \qquad \text{if } \mathbf{S}\left(t_m^{\star}\right) < u_i \leq \mathbf{S}\left(t_{m-1}^{\star}\right)$$

Method of inversion

Piecewise distribution functions

Example #6

We model the default time $\boldsymbol{\tau}$ with the piecewise exponential model and the following parameters:

$$\lambda = \begin{cases} 5\% & \text{if} \quad t \text{ is less or equal than one year} \\ 8\% & \text{if} \quad t \text{ is between one and five years} \\ 12\% & \text{if} \quad t \text{ is larger than five years} \end{cases}$$

Method of inversion
Piecewise distribution functions

We have $\mathbf{S}(0) = 1$, $\mathbf{S}(1) = 0.9512$ and $\mathbf{S}(5) = 0.6907$. We deduce that:

$$t_i \leftarrow \begin{cases} 0 + (1/0.05) \cdot \ln(1/u_i) & \text{if} \quad u_i \in [0.9512, 1] \\ 1 + (1/0.08) \cdot \ln(0.9512/u_i) & \text{if} \quad u_i \in [0.6907, 0.9512[\\ 5 + (1/0.12) \cdot \ln(0.6907/u_i) & \text{if} \quad u_i \in [0, 0.6907[\end{cases}$$

Table: Simulation of the piecewise exponential model

u_i	$t^\star_{m-1}$	$S\left(t^\star_{m-1}\right)$	λ_m	t_i
0.9950	0	1.0000	0.05	0.1003
0.3035	5	0.6907	0.12	11.8531
0.5429	5	0.6907	0.12	7.0069
0.9140	1	0.9512	0.08	1.4991
0.7127	1	0.9512	0.08	4.6087

Method of transformation

Let $\{Y_1, Y_2, \ldots\}$ be a vector of independent random variables. The simulation of the random variable $X = g(Y_1, Y_2, \ldots)$ is straightforward if we know how to easily simulate the random variables Y_i. We notice that the inversion method is a particular case of the transform method, because we have:

$$X = g(U) = \mathbf{F}^{-1}(U)$$

Method of transformation

- The Binomial random variable is the sum of n *iid* Bernoulli random variables:

$$\mathcal{B}(n,p)=\sum_{i=1}^{n}\mathcal{B}_i(p)$$

- We simulate the Binomial random variate x using n uniform random numbers:

$$x=\sum_{i=1}^{n}\mathbb{1}\left\{u_i\leq p\right\}$$

Method of transformation

To simulate the chi-squared random variable $\chi^2(\nu)$, we can use the following relationship:

$$\chi^2(\nu) = \sum_{i=1}^{\nu} \chi_i^2(1) = \sum_{i=1}^{\nu} (\mathcal{N}_i(0,1))^2$$

Method of transformation

Box-Muller algorithm

if U_1 and U_2 are two independent uniform random variables, then X_1 and X_2 defined by:

$$\left\{ \begin{array}{l} X_1 = \sqrt{-2 \ln U_1} \cdot \cos\left(2\pi U_2\right) \\ X_2 = \sqrt{-2 \ln U_1} \cdot \sin\left(2\pi U_2\right) \end{array} \right.$$

are independent and follow the Gaussian distribution distribution $\mathcal{N}\left(0,1\right)$

Method of transformation

- If N_t is a Poisson process with intensity λ, the duration T between two consecutive events is an exponential:

$$\Pr(T \leq t) = 1 - e^{-\lambda t}$$

- Since the durations are independent, we have:

$$T_1 + T_2 + \ldots + T_n = \sum_{i=1}^{n} E_i$$

where $E_i \sim \mathcal{E}(\lambda)$

- Because the Poisson random variable is the number of events that occur in the unit interval of time, we also have:

$$X = \max\{n : T_1 + T_2 + \ldots + T_n \leq 1\} = \max\left\{n : \sum_{i=1}^{n} E_i \leq 1\right\}$$

Method of transformation

- We notice that:

$$\sum_{i=1}^{n} E_i = -\frac{1}{\lambda}\sum_{i=1}^{n} \ln U_i = -\frac{1}{\lambda}\ln \prod_{i=1}^{n} U_i$$

  where U_i are *iid* uniform random variables

- We deduce that:

$$X = \max \left\{ n : -\frac{1}{\lambda}\ln \prod_{i=1}^{n} U_i \leq 1 \right\} = \max \left\{ n : \prod_{i=1}^{n} U_i \geq e^{-\lambda} \right\}$$

Method of transformation

We can then simulate the Poisson random variable with the following algorithm:

1. set $n = 0$ and $p = 1$;
2. calculate $n = n + 1$ and $p = p \cdot u_i$ where u_i is a uniform random variate;
3. if $p \geq e^{-\lambda}$, go back to step 2; otherwise, return $X = n - 1$

Rejection sampling

Theorem

- $\mathbf{F}(x)$ and $\mathbf{G}(x)$ are two distribution functions such that $f(x) \leq cg(x)$ for all x with $c > 1$
- We note $X \sim \mathbf{G}$ and consider an independent uniform random variable $U \sim \mathcal{U}_{[0,1]}$
- Then, the conditional distribution function of X given that $U \leq f(X) / (cg(X))$ is $\mathbf{F}(x)$

Rejection sampling

Proof

Let us introduce the random variables B and Z:

$$B = \mathbb{1}\left\{U \leq \frac{f(X)}{cg(X)}\right\} \qquad \text{and} \qquad Z = X \left| U \leq \frac{f(X)}{cg(X)} \right.$$

We have:

$$\begin{aligned}
\Pr\{B = 1\} &= \Pr\left\{U \leq \frac{f(X)}{cg(X)}\right\} \\
&= \mathbb{E}\left[\frac{f(X)}{cg(X)}\right] = \int_{-\infty}^{+\infty} \frac{f(x)}{cg(x)} g(x)\,\mathrm{d}x \\
&= \frac{1}{c}\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x \\
&= \frac{1}{c}
\end{aligned}$$

Rejection sampling

Proof

The distribution function of Z is defined by:

$$\Pr\{Z \leq x\} = \Pr\left\{X \leq x \,\middle|\, U \leq \frac{f(X)}{cg(X)}\right\}$$

We deduce that:

$$\begin{aligned}\Pr\{Z \leq x\} &= \frac{\Pr\left\{X \leq x, U \leq \dfrac{f(X)}{cg(X)}\right\}}{\Pr\left\{U \leq \dfrac{f(X)}{cg(X)}\right\}} = c\int_{-\infty}^{x}\int_{0}^{f(x)/(cg(x))} g(x)\,\mathrm{d}u\,\mathrm{d}x \\ &= c\int_{-\infty}^{x}\frac{f(x)}{cg(x)}g(x)\,\mathrm{d}x = \int_{-\infty}^{x} f(x)\,\mathrm{d}x \\ &= \mathbf{F}(x)\end{aligned}$$

This proves that $Z \sim \mathbf{F}$

Rejection sampling

Acceptance-rejection algorithm

1. generate two independent random variates x and u from **G** and $\mathcal{U}_{[0,1]}$;
2. calculate v as follows:
$$v = \frac{f(x)}{cg(x)}$$
3. if $u \leq v$, return x ('*accept*'); otherwise, go back to step 1 ('*reject*')

Remark

The underlying idea of this algorithm is then to simulate the distribution function **F** by assuming that it is easier to generate random numbers from **G**, which is called the proposal distribution. However, some of these random numbers must be '*rejected*', because the function $c \cdot g(x)$ '*dominates*' the density function $f(x)$

Rejection sampling

- The number of iterations N needed to successfully generate Z has a geometric distribution $\mathcal{G}(p)$, where $p = \Pr\{B = 1\} = c^{-1}$ is the acceptance ratio
- The average number of iterations is equal to:

$$\mathbb{E}[N] = 1/p = c$$

- To maximize the efficiency (or the acceptance ratio) of the algorithm, we have to choose the constant c such that:

$$c = \sup_x \frac{f(x)}{g(x)}$$

Rejection sampling

- We consider the normal distribution $\mathcal{N}(0,1)$
- We use the Cauchy distribution function as the proposal distribution:

$$g(x) = \frac{1}{\pi(1+x^2)}$$

- We can show that:

$$\phi(x) \leq \frac{\sqrt{2\pi}}{e^{0.5}} g(x)$$

  meaning that $c \approx 1.52$
- We have:

$$\mathbf{G}(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x$$

  and:

$$\mathbf{G}^{-1}(u) = \tan\left(\pi\left(u - \frac{1}{2}\right)\right)$$

Rejection sampling

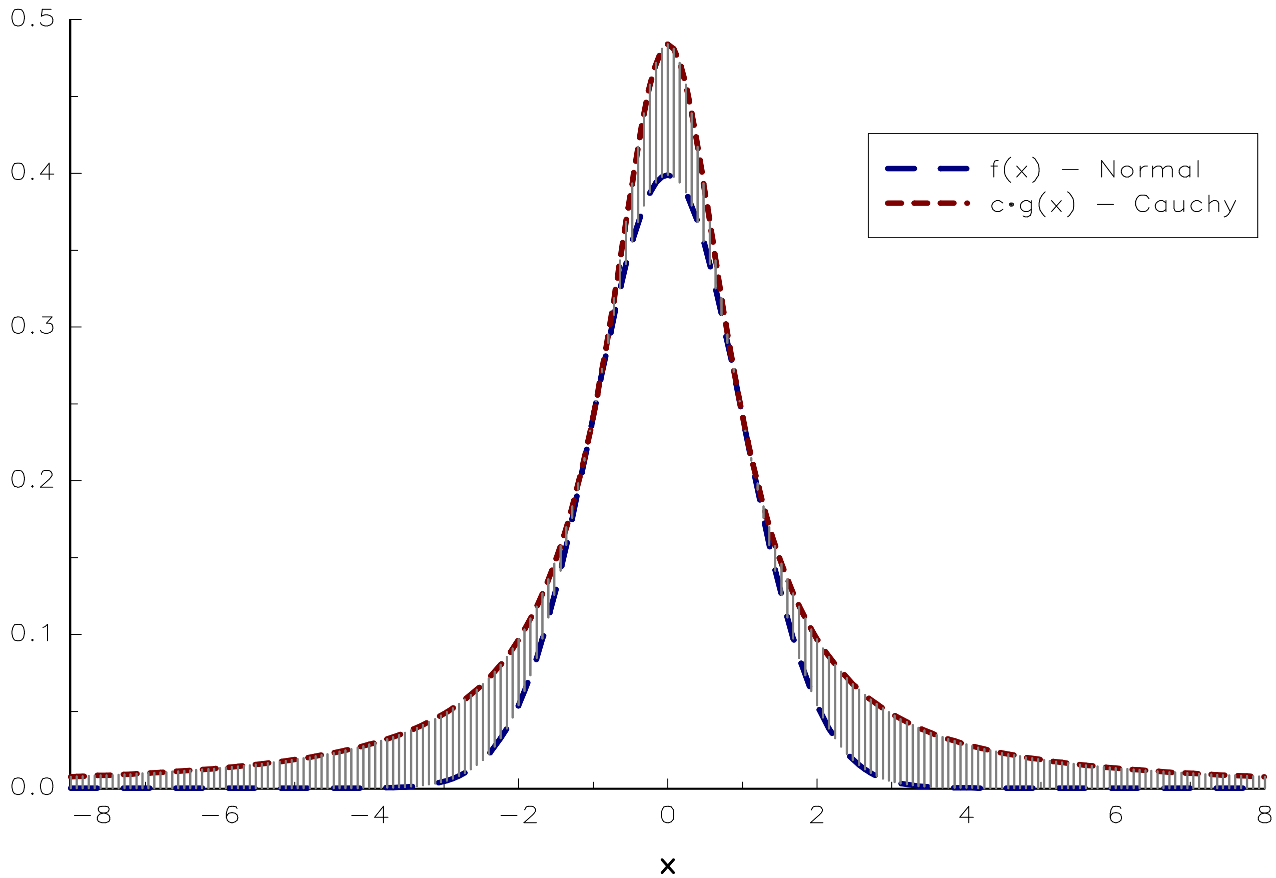


Figure: Rejection sampling applied to the normal distribution

Rejection sampling

Acceptance-rejection algorithm for simulating $\mathcal{N}(0,1)$

1. generate two independent uniform random variates u_1 and u_2 and set:
$$x \leftarrow \tan\left(\pi\left(u_1 - \frac{1}{2}\right)\right)$$
2. calculate v as follows:
$$v = \frac{e^{0.5}\phi(x)}{\sqrt{2\pi}g(x)} = \frac{(1+x^2)}{2e^{(x^2-1)/2}}$$
3. if $u_2 \leq v$, accept x; otherwise, go back to step 1

The acceptance ratio is $1/1.52 \approx 65.8\%$

Rejection sampling

Table: Simulation of the standard Gaussian distribution using the acceptance-rejection algorithm

u_1	u_2	x	v	test	z
0.9662	0.1291	9.3820	0.0000	reject	
0.0106	0.1106	−30.0181	0.0000	reject	
0.3120	0.8253	−0.6705	0.9544	accept	−0.6705
0.9401	0.9224	5.2511	0.0000	reject	
0.2170	0.4461	−1.2323	0.9717	accept	−1.2323
0.6324	0.0676	0.4417	0.8936	accept	0.4417
0.6577	0.1344	0.5404	0.9204	accept	0.5404
0.1596	0.6670	−1.8244	0.6756	accept	−1.8244
0.4183	0.3872	−0.2625	0.8513	accept	−0.2625
0.9625	0.0752	8.4490	0.0000	reject	

Rejection sampling

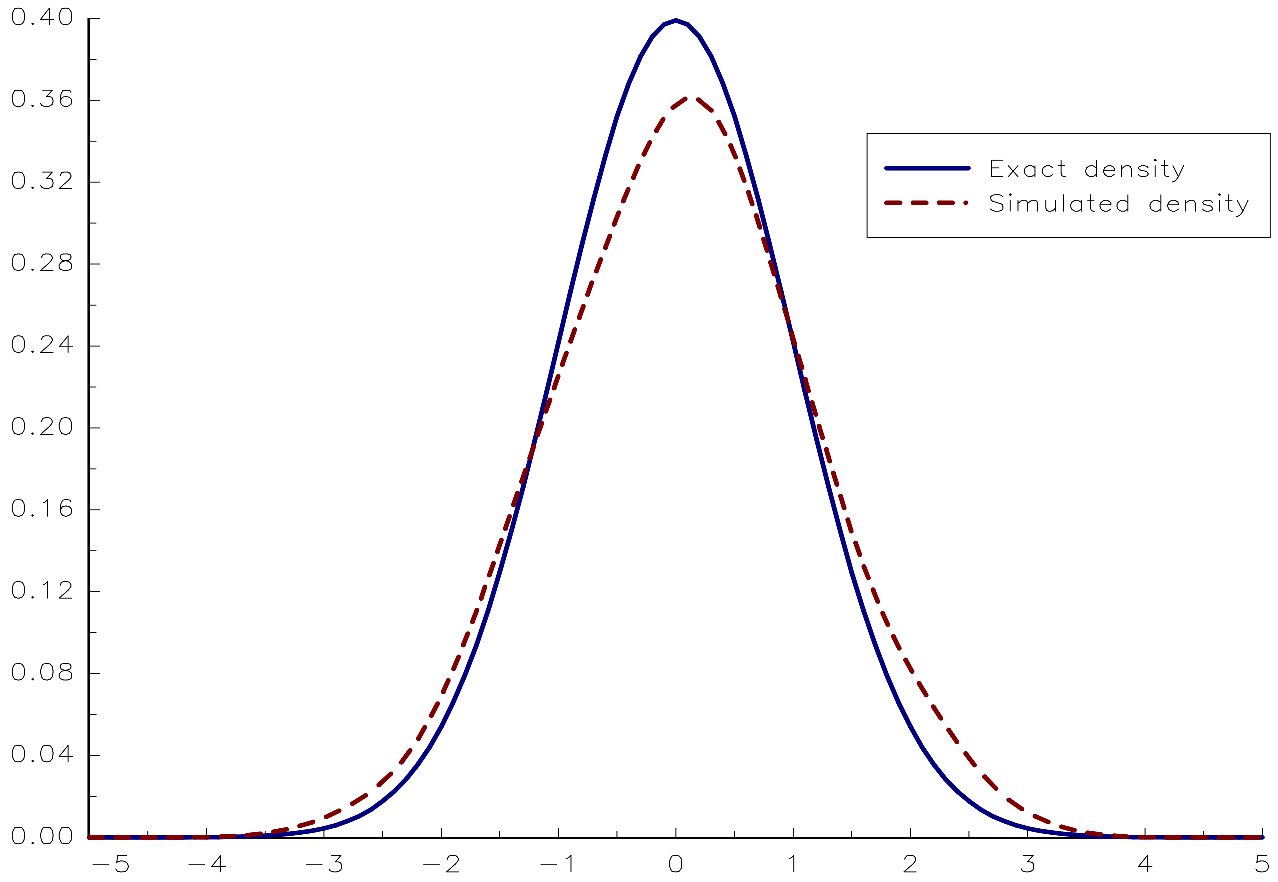


Figure: Comparison of the exact and simulated densities

Method of mixtures

- A finite mixture can be decomposed as a weighted sum of distribution functions:

$$\mathbf{F}(x) = \sum_{k=1}^{n} \pi_k \cdot \mathbf{G}_k(x)$$

  where $\pi_k \geq 0$ and $\sum_{k=1}^{n} \pi_k = 1$
- The probability density function is:

$$f(x) = \sum_{k=1}^{n} \pi_k \cdot g_k(x)$$

- To simulate the probability distribution $\mathbf{F}$, we introduce the random variable B, whose probability mass function is defined by:

$$p(k) = \Pr\{B = k\} = \pi_k$$

  It follows that:

$$\mathbf{F}(x) = \sum_{k=1}^{n} \Pr\{B = k\} \cdot \mathbf{G}_k(x)$$

Method of mixtures

We deduce the following algorithm:

1. generate the random variate b from the probability mass function $p(k)$
2. generate the random variate x from the probability distribution $\mathbf{G}_b(x)$

Method of mixtures

The previous approach can be easily extended to continuous mixtures:

$$f(x) = \int_{\Omega} \pi(\omega) \, g(x;\omega) \, \mathrm{d}\omega$$

where $\omega \in \Omega$ is a parameter of the distribution $\mathbf{G}$

Method of mixtures

The negative binomial distribution is a gamma-Poisson mixture distribution:

$$\left\{ \begin{array}{l} \mathcal{NB}\left(r,p\right) \sim \mathcal{P}\left(\Lambda\right) \\ \Lambda \sim \mathcal{G}\left(r,\left(1-p\right)/p\right) \end{array} \right.$$

To simulate the negative binomial distribution, we simulate

1. the gamma random variate $g \sim \mathcal{G}\left(r,\left(1-p\right)/p\right)$
2. and then the Poisson random variable p, whose parameter λ is equal to g

Random vectors

The random vector $X = (X_1, \ldots, X_n)$ has a given distribution function $\mathbf{F}(x) = \mathbf{F}(x_1, \ldots, x_n)$

Method of conditional distributions

- If $X_1, \ldots, X_n$ are independent, we have:

$$\mathbf{F}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \mathbf{F}_i(x_i)$$

To simulate X, we can then generate each component $X_i \sim \mathbf{F}_i$ individually, for example by applying the method of inversion

Method of conditional distributions

- If $X_1, \ldots, X_n$ are dependent, we have:

$$\begin{aligned}\mathbf{F}(x_1, \ldots, x_n) &= \mathbf{F}_1(x_1)\,\mathbf{F}_{2|1}(x_2 \mid x_1)\,\mathbf{F}_{3|1,2}(x_3 \mid x_1, x_2) \times \cdots \times \\ &\quad \mathbf{F}_{n|1,\ldots,n-1}(x_n \mid x_1, \ldots, x_{n-1}) \\ &= \prod_{i=1}^{n} \mathbf{F}_{i|1,\ldots,i-1}(x_i \mid x_1, \ldots, x_{i-1})\end{aligned}$$

  where $\mathbf{F}_{i|1,\ldots,i-1}(x_i \mid x_1, \ldots, x_{i-1})$ is the conditional distribution of X_i given $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}$
- This '*conditional*' random variable is denoted by $Y_i = X_i \mid X_1 = x_1, \ldots, X_{i-1} = x_{i-1}$
- The random variables $(Y_1, \ldots, Y_n)$ are independent

Method of conditional distributions

We obtain the following algorithm:

1. generate x_1 from $\mathbf{F}_1(x)$ and set $i = 2$
2. generate x_i from $\mathbf{F}_{i|1,\ldots,i-1}(x \mid x_1, \ldots, x_{i-1})$ given $X_1 = x_1, \ldots, X_{i-1} = x_{i-1}$ and set $i = i + 1$
3. repeat step 2 until $i = n$

Method of conditional distributions

$\mathbf{F}_{i|1,\ldots,i-1}\left(x \mid x_1,\ldots,x_{i-1}\right)$ is a univariate distribution function, which depends on the argument x and parameters $x_1,\ldots,x_{i-1}$. To simulate it, we can therefore use the method of inversion:

$$x_i \leftarrow \mathbf{F}_{i|1,\ldots,i-1}^{-1}\left(u_i \mid x_1,\ldots,x_{i-1}\right)$$

where $\mathbf{F}_{i|1,\ldots,i-1}^{-1}$ is the inverse of the conditional distribution function and u_i is a uniform random variate

Method of conditional distributions

Example #7

We consider the bivariate logistic distribution defined as:

$$\mathbf{F}(x_1, x_2) = \left(1 + e^{-x_1} + e^{-x_2}\right)^{-1}$$

Method of conditional distributions

We have $\mathbf{F}_1(x_1) = \mathbf{F}(x_1, +\infty) = \left(1 + e^{-x_1}\right)^{-1}$. We deduce that the conditional distribution of X_2 given $X_1 = x_1$ is:

$$\begin{aligned}\mathbf{F}_{2|1}(x_2 \mid x_1) &= \frac{\mathbf{F}(x_1, x_2)}{\mathbf{F}_1(x_1)} \\ &= \frac{1 + e^{-x_1}}{1 + e^{-x_1} + e^{-x_2}}\end{aligned}$$

We obtain:

$$\mathbf{F}_1^{-1}(u) = \ln u - \ln(1 - u)$$

and:

$$\mathbf{F}_{2|1}^{-1}(u \mid x_1) = \ln u - \ln(1 - u) - \ln\left(1 + e^{-x_1}\right)$$

Method of conditional distributions

We deduce the following algorithm:

1. generate two independent uniform random variates u_1 and u_2;
2. generate x_1 from u_1:

$$x_1 \leftarrow \ln u_1 - \ln (1 - u_1)$$

3. generate x_2 from u_2 and x_1:

$$x_2 \leftarrow \ln u_2 - \ln (1 - u_2) - \ln \left(1 + e^{-x_1}\right)$$

Because we have $(1 + e^{-x_1})^{-1} = u_1$, the last step can be replaced by:

3. generate x_2 from u_2 and u_1:

$$x_2 \leftarrow \ln \left(\frac{u_1 u_2}{1 - u_2}\right)$$

Method of conditional distributions

- The method of conditional distributions can be used for simulating uniform random vectors $(U_1, \ldots, U_n)$ generated by copula functions
- We have

$$\begin{aligned} \mathbf{C}(u_1, \ldots, u_n) &= \mathbf{C}_1(u_1)\,\mathbf{C}_{2|1}(u_2 \mid u_1)\,\mathbf{C}_{3|1,2}(u_3 \mid u_1, u_2) \times \cdots \times \\ &\quad\ \mathbf{C}_{n|1,\ldots,n-1}(u_n \mid u_1, \ldots, u_{n-1}) \\ &= \prod_{i=1}^{n} \mathbf{C}_{i|1,\ldots,i-1}(u_i \mid u_1, \ldots, u_{i-1}) \end{aligned}$$

  where $\mathbf{C}_{i|1,\ldots,i-1}(u_i \mid u_1, \ldots, u_{i-1})$ is the conditional distribution of U_i given $U_1 = u_1, \ldots, U_{i-1} = u_{i-1}$
- By definition, we have $\mathbf{C}_1(u_1) = u_1$

Method of conditional distributions

We obtain the following algorithm:

1. generate n independent uniform random variates $v_1, \ldots, v_n$;
2. generate $u_1 \leftarrow v_1$ and set $i = 2$;
3. generate u_i by finding the root of the equation:

$$\mathbf{C}_{i|1,\ldots,i-1}\left(u_i \mid u_1, \ldots, u_{i-1}\right) = v_i$$

   and set $i = i + 1$;
4. repeat step 3 until $i = n$.

For some copula functions, there exists an analytical expression of the inverse of the conditional copula. In this case, the third step is replaced by:

3. generate u_i by the inversion method:

$$u_i \leftarrow \mathbf{C}_{i|1,\ldots,i-1}^{-1}\left(v_i \mid u_1, \ldots, u_{i-1}\right)$$

Method of conditional distributions

For any probability distribution, the conditional distribution can be calculated as follows:

$$\mathbf{F}_{i|1,\ldots,i-1}\left(x_i \mid x_1,\ldots,x_{i-1}\right) = \frac{\mathbf{F}\left(x_1,\ldots,x_{i-1},x_i\right)}{\mathbf{F}\left(x_1,\ldots,x_{i-1}\right)}$$

In particular, we have:

$$\begin{array}{rcl} \partial_1 \mathbf{F}\left(x_1,x_2\right) & = & \partial_1 \left(\mathbf{F}_1\left(x_1\right)\cdot \mathbf{F}_{2|1}\left(x_2 \mid x_1\right)\right) \\ & = & f_1\left(x_1\right)\cdot \mathbf{F}_{2|1}\left(x_2 \mid x_1\right) \end{array}$$

For copula functions, the density $f_1\left(x_1\right)$ is equal to 1, meaning that:

$$\mathbf{C}_{2|1}\left(u_2 \mid u_1\right) = \partial_1 \mathbf{C}\left(u_1,u_2\right)$$

We can generalize this result and show that the conditional copula given some random variables U_i for $i \in \Omega$ is equal to the cross-derivative of the copula function $\mathbf{C}$ with respect to the arguments u_i for $i \in \Omega$

Method of conditional distributions

- Archimedean copulas are defined as:

$$\mathbf{C}(u_1, u_2) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2))$$

  where $\varphi(u)$ is the generator function
- We have:

$$\varphi(\mathbf{C}(u_1, u_2)) = \varphi(u_1) + \varphi(u_2)$$

  and:

$$\varphi'(\mathbf{C}(u_1, u_2)) \cdot \frac{\partial\, \mathbf{C}(u_1, u_2)}{\partial\, u_1} = \varphi'(u_1)$$

- We deduce the following expression of the conditional copula:

$$\mathbf{C}_{2|1}(u_2 \mid u_1) = \frac{\partial\, \mathbf{C}(u_1, u_2)}{\partial\, u_1} = \frac{\varphi'(u_1)}{\varphi'(\varphi^{-1}(\varphi(u_1) + \varphi(u_2)))}$$

- The calculation of the inverse function gives:

$$\mathbf{C}_{2|1}^{-1}(v \mid u_1) = \varphi^{-1}\left(\varphi\left(\varphi'^{-1}\left(\frac{\varphi'(u_1)}{v}\right)\right) - \varphi(u_1)\right)$$

Method of conditional distributions

We obtain the following algorithm for simulating Archimedean copulas:

1. generate two independent uniform random variates v_1 and v_2;
2. generate $u_1 \leftarrow v_1$;
3. generate u_2 by the inversion method:

$$u_2 \leftarrow \varphi^{-1}\left(\varphi\left(\varphi'^{-1}\left(\frac{\varphi'(u_1)}{v_2}\right)\right) - \varphi(u_1)\right)$$

Method of conditional distributions

Example #8

We consider the Clayton copula:

$$\mathbf{C}(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$$

Method of conditional distributions

The Clayton copula is an Archimedean copula, whose generator function is:

$$\varphi(u) = u^{-\theta} - 1$$

We deduce that:

$$\begin{array}{rcl} \varphi^{-1}(u) & = & (1+u)^{-1/\theta} \\ \varphi'(u) & = & -\theta u^{-(\theta+1)} \\ \varphi'^{-1}(u) & = & (-u/\theta)^{-1/(\theta+1)} \end{array}$$

We obtain:

$$\mathbf{C}_{2|1}^{-1}(v \mid u_1) = \left(1 + u_1^{-\theta}\left(v^{-\theta/(\theta+1)} - 1\right)\right)^{-1//\theta}$$

Method of conditional distributions

Table: Simulation of the Clayton copula

Random uniform variates		Clayton copula $\theta = 0.01$		Clayton copula $\theta = 1.5$	
v_1	v_2	u_1	u_2	u_1	u_2
0.2837	0.4351	0.2837	0.4342	0.2837	0.3296
0.0386	0.2208	0.0386	0.2134	0.0386	0.0297
0.3594	0.5902	0.3594	0.5901	0.3594	0.5123
0.3612	0.3268	0.3612	0.3267	0.3612	0.3247
0.0797	0.6479	0.0797	0.6436	0.0797	0.1704

Method of transformation

- To simulate a Gaussian random vector $X \sim \mathcal{N}(\mu, \Sigma)$, we consider the following transformation:

$$X = \mu + A \cdot N$$

  where $AA^\top = \Sigma$ and $N \sim \mathcal{N}(\mathbf{0}, I)$
- Since Σ is a positive definite symmetric matrix, it has a unique Cholesky decomposition:

$$\Sigma = PP^\top$$

  where P is a lower triangular matrix

Method of transformation

The decomposition $AA^\top = \Sigma$ is not unique. For instance, if we use the eigendecomposition:

$$\Sigma = U \Lambda U^\top$$

we can set $A = U \Lambda^{1/2}$. Indeed, we have:

$$\begin{aligned} AA^\top &= U \Lambda^{1/2} \Lambda^{1/2} U^\top \\ &= U \Lambda U^\top \\ &= \Sigma \end{aligned}$$

Method of transformation

To simulate a multivariate Student's t distribution $Y = (Y_1, \ldots, Y_n) \sim \mathbf{T}_n(\Sigma, \nu)$, we use the relationship:

$$Y_i = \frac{X_i}{\sqrt{Z/\nu}}$$

where the random vector $X = (X_1, \ldots, X_n) \sim \mathcal{N}(\mathbf{0}, \Sigma)$ and the random variable $Z \sim \chi^2(\nu)$ are independent

Method of transformation

- If $X = (X_1, \ldots, X_n) \sim \mathbf{F}$, then the probability distribution of the random vector $U = (U_1, \ldots, U_n)$ defined by:

$$U_i = \mathbf{F}_i(X)$$

  is the copula function $\mathbf{C}$ associated to $\mathbf{F}$
- To simulate the Normal copula with the matrix of parameters ρ, we simulate $N \sim \mathcal{N}(\mathbf{0}, I)$ and apply the transformation:

$$U = \Phi(P \cdot N)$$

  where P is the Cholesky decomposition of the correlation matrix ρ
- To simulate the Student's t copulawith the matrix of parameters ρ and ν degrees of freedom, we simulate $T \sim \mathbf{T}_n(\rho, \nu)$ and apply the transformation:

$$U_i = \mathbf{T}_v(T_i)$$

Method of transformation

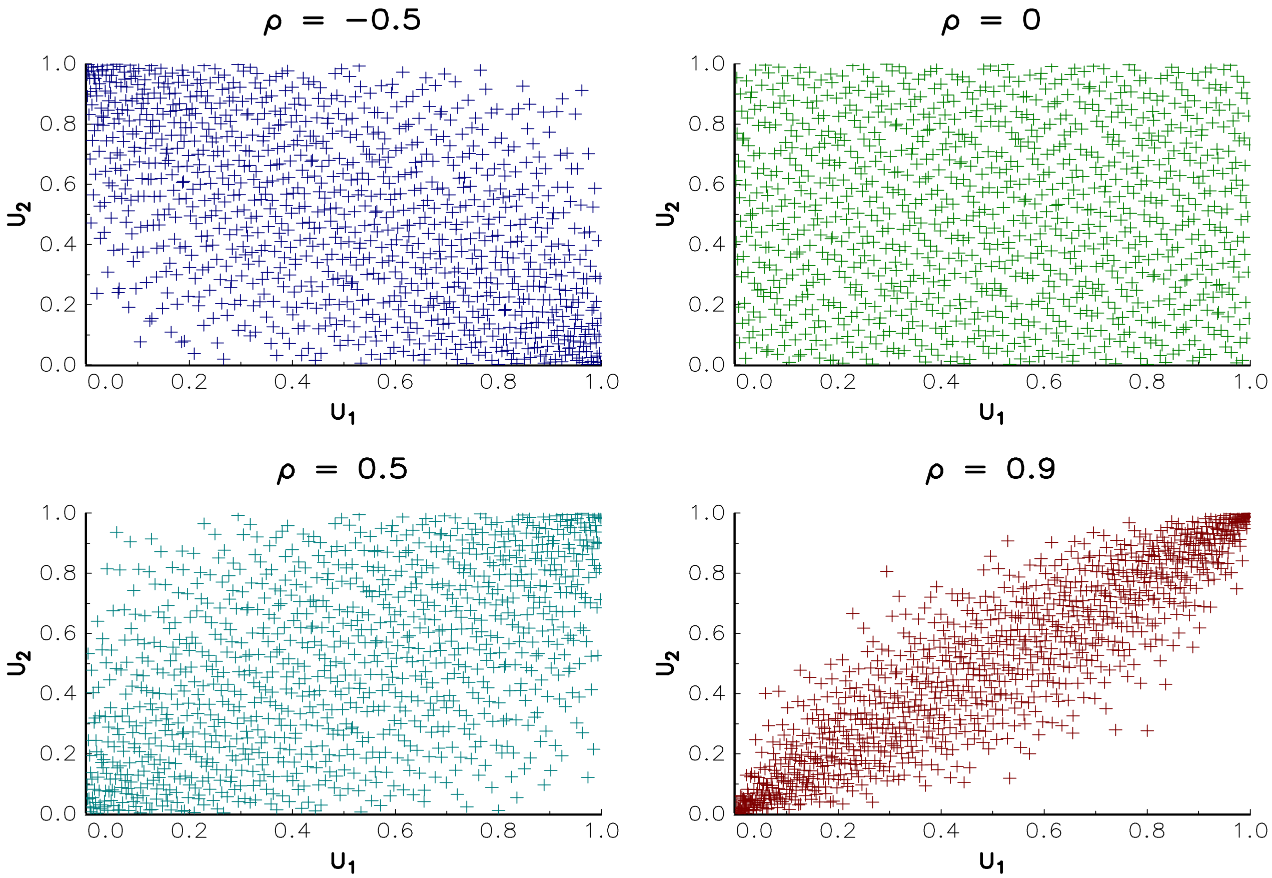


Figure: Simulation of the Normal copula

Method of transformation

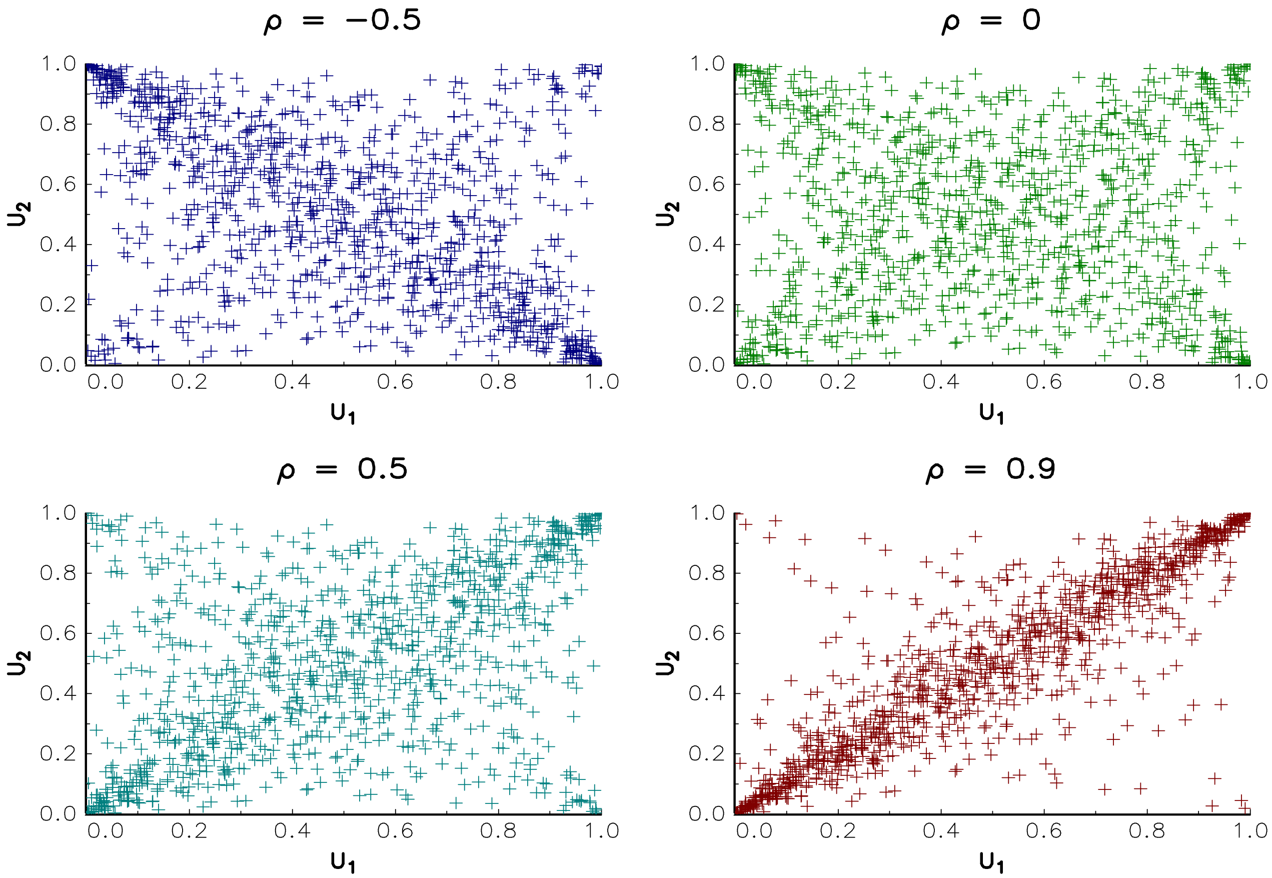


Figure: Simulation of the t_1 copula

Method of transformation

Frailty copulas are defined as:

$$\mathbf{C}(u_1, \ldots, u_n) = \psi\left(\psi^{-1}(u_1) + \ldots + \psi^{-1}(u_n)\right)$$

where $\psi(x)$ is the Laplace transform of a random variable X
They can be generated using the following algorithm:

1. simulate n independent uniform random variates $v_1, \ldots, v_n$;
2. simulate the frailty random variate x with the Laplace transform ψ;
3. apply the transformation:

$$(u_1, \ldots, u_n) \leftarrow \left(\psi\left(-\frac{\ln u_1}{x}\right), \ldots, \psi\left(-\frac{\ln u_n}{x}\right)\right)$$

Method of transformation

- The Clayton copula is a frailty copula where $\psi(x) = (1+x)^{-1/\theta}$ is the Laplace transform of the gamma random variable $\mathcal{G}(1/\theta, 1)$
- The algorithm to simulate the Clayton copula is:

$$\left\{ \begin{array}{l} x \leftarrow \mathcal{G}(1/\theta, 1) \\ (u_1, \ldots, u_n) \leftarrow \left(\left(1 - \dfrac{\ln u_1}{x}\right)^{-1/\theta}, \ldots, \left(1 - \dfrac{\ln u_n}{x}\right)^{-1/\theta} \right) \end{array} \right.$$

Method of transformation

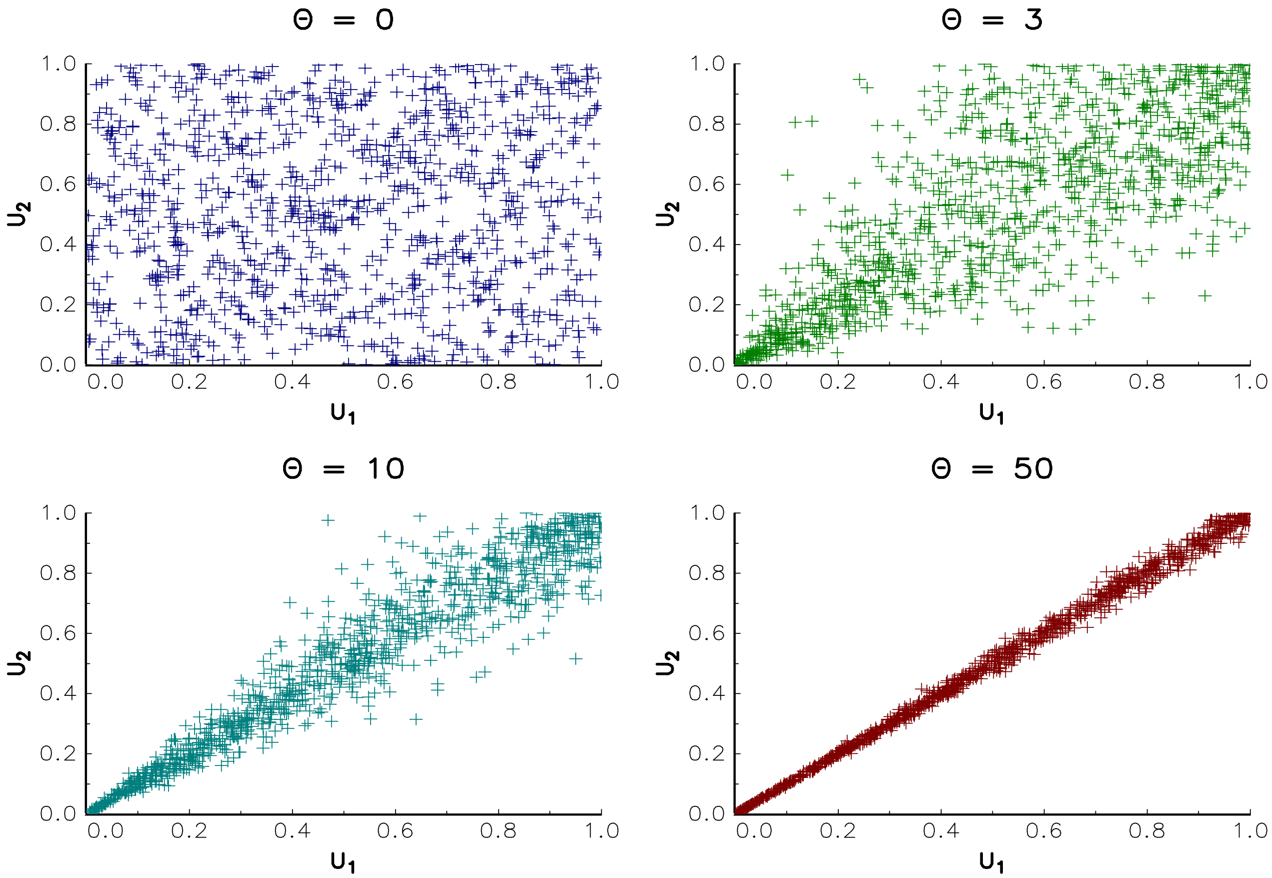


Figure: Simulation of the Clayton copula

Method of transformation

- We consider the multivariate distribution $\mathbf{F}(x_1, \ldots, x_n)$, whose canonical decomposition is defined as:

$$\mathbf{F}(x_1, \ldots, x_n) = \mathbf{C}(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_n(x_n))$$

- If $(U_1, \ldots, U_n) \sim \mathbf{C}$, the random vector $(X_1, \ldots, X_n) = \left(\mathbf{F}_1^{-1}(U_1), \ldots, \mathbf{F}_n^{-1}(U_n)\right)$ follows the distribution function $\mathbf{F}$
- We deduce the following algorithm:

$$\left\{ \begin{array}{l} (u_1, \ldots, u_n) \leftarrow \mathbf{C} \\ (x_1, \ldots, x_n) \leftarrow \left(\mathbf{F}_1^{-1}(u_1), \ldots, \mathbf{F}_n^{-1}(u_n)\right) \end{array} \right.$$

Method of transformation

- We assume that $\boldsymbol{\tau} \sim \mathcal{E}(5\%)$ and $\mathrm{LGD} \sim \mathcal{B}(2,2)$
- We also assume that the default time and the loss given default are correlated and the dependence function is a Clayton copula

Method of transformation

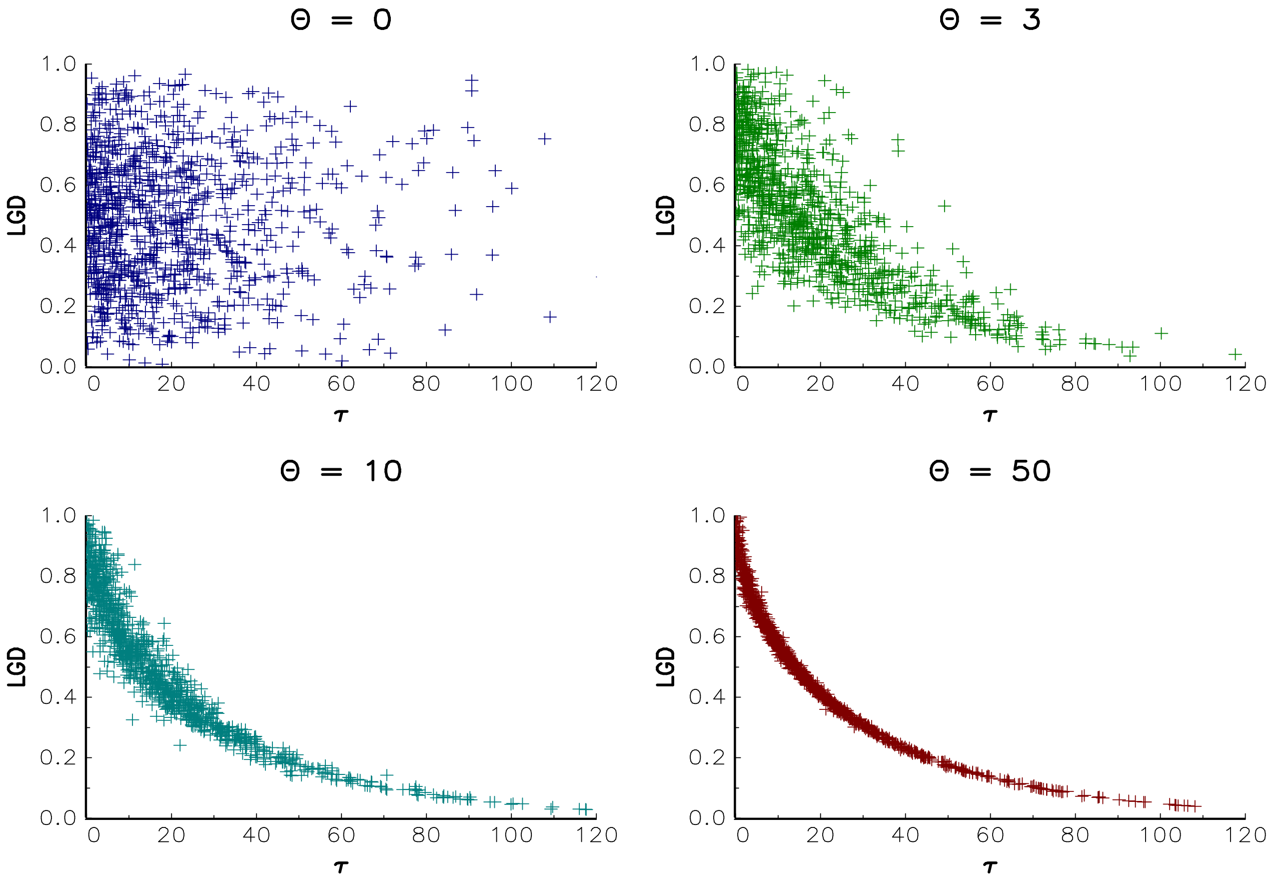


Figure: Simulation of the correlated random vector (τ, LGD)

Method of transformation

Remark

The previous algorithms suppose that we know the analytical expression $\mathbf{F}_i$ of the univariate probability distributions in order to calculate the quantile $\mathbf{F}_i^{-1}$. This is not always the case. For instance, in the operational risk, the loss of the bank is equal to the sum of aggregate losses:

$$L = \sum_{k=1}^{K} S_k$$

where S_k is also the sum of individual losses for the k^{th} cell of the mapping matrix. In practice, the probability distribution of S_k is estimated by the method of simulations

Method of transformation

The method of the empirical quantile function is implemented as follows:

1. for each random variable X_i, simulate m_1 random variates $x_{i,m}^{\star}$ and estimate the empirical distribution $\hat{\mathbf{F}}_i$;
2. simulate a random vector $(u_1, \ldots, u_n)$ from the copula function $\mathbf{C}(u_1, \ldots, u_n)$;
3. simulate the random vector $(x_1, \ldots, x_n)$ by inverting the empirical distributions $\hat{\mathbf{F}}_i$:
$$x_i \leftarrow \hat{\mathbf{F}}_i^{-1}(u_i)$$
we also have:
$$x_i \leftarrow \inf \left\{ x \,\middle|\, \frac{1}{m_1} \sum_{m=1}^{m_1} \mathbf{1}\left\{x \leq x_{i,m}^{\star}\right\} \geq u_i \right\}$$
4. repeat steps 2 and 3 m_2 times

Method of transformation

- $X_1 \sim \mathcal{N}(0,1)$
- $X_2 \sim \mathcal{N}(0,1)$
- The dependence function of (X_1, X_2) is the Clayton copula with parameter $\theta = 3$

Method of transformation

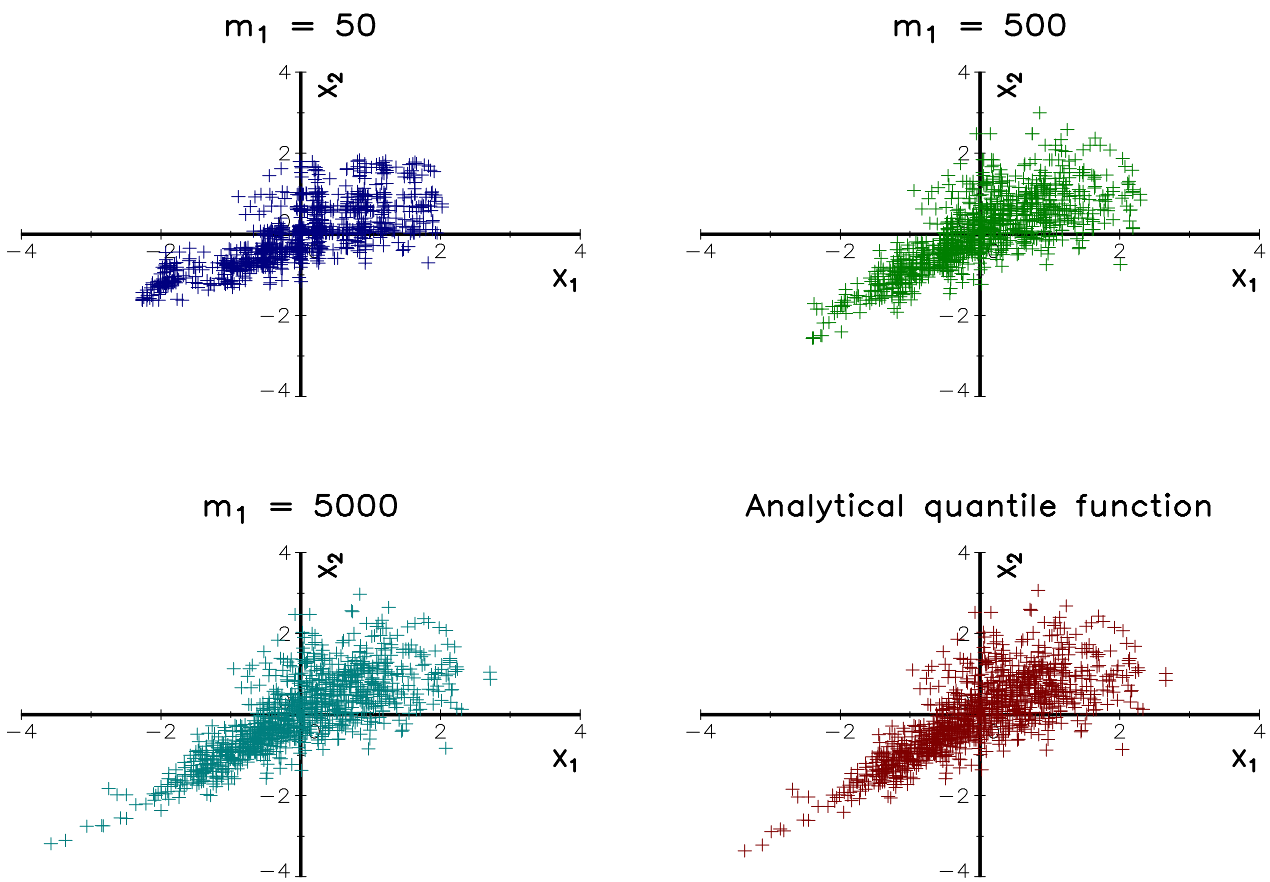


Figure: Convergence of the method of the empirical quantile function

Method of transformation

- $X_1 \sim \mathcal{N}(-1, 2)$, $X_2 \sim \mathcal{N}(0, 1)$, $Y_1 \sim \mathcal{G}(0.5)$ and $Y_2 \sim \mathcal{G}(1, 2)$ are four independent random variables
- Let $(Z_1 = X_1 + Y_1, Z_2 = X_2 \cdot Y_2)$ be the random vector
- The dependence function of Z is the t copula with parameters $\nu = 2$ and $\rho = -70\%$
- It is not possible to find an analytical expression of the marginal distributions of Z_1 and Z_2

Method of transformation

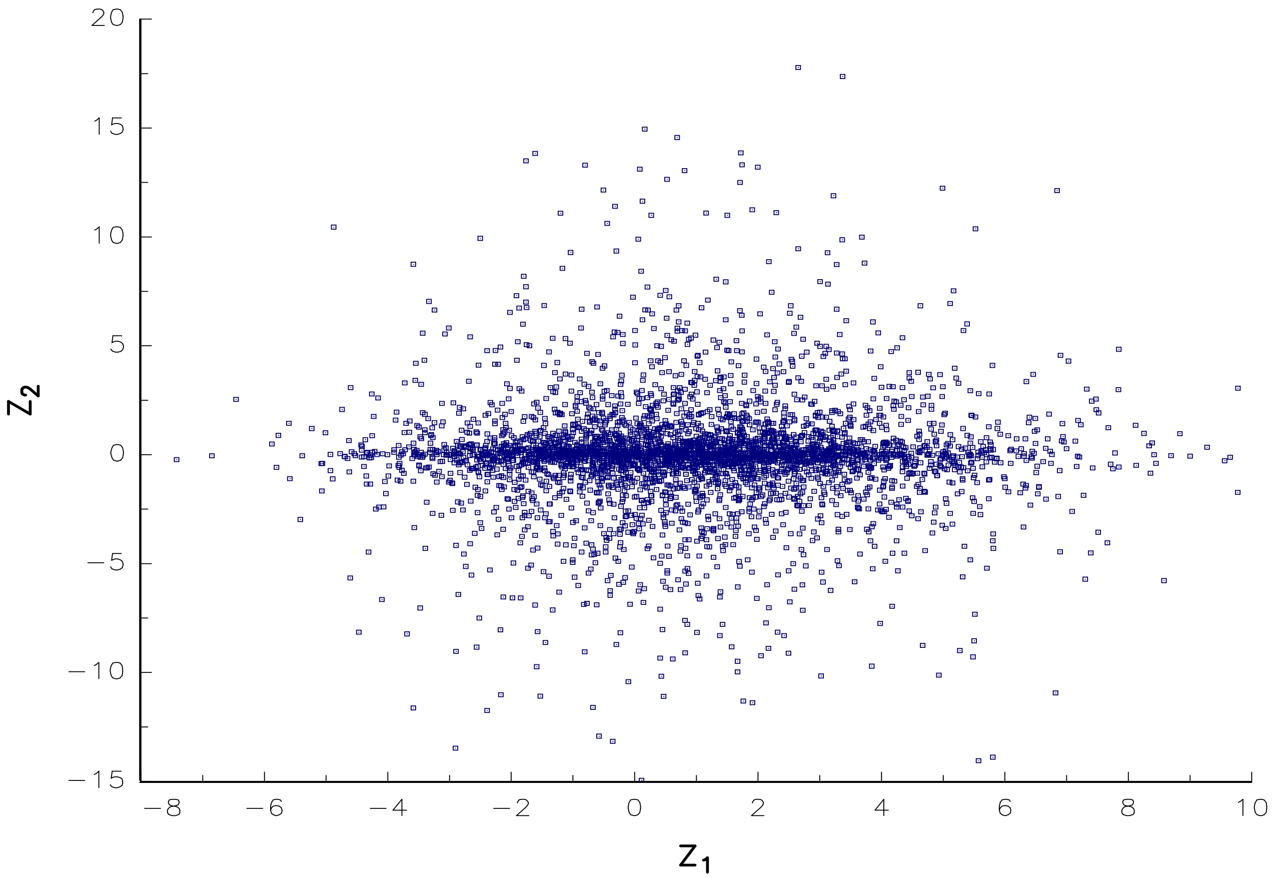


Figure: Simulation of the random variables Z_1 and Z_2

Method of transformation

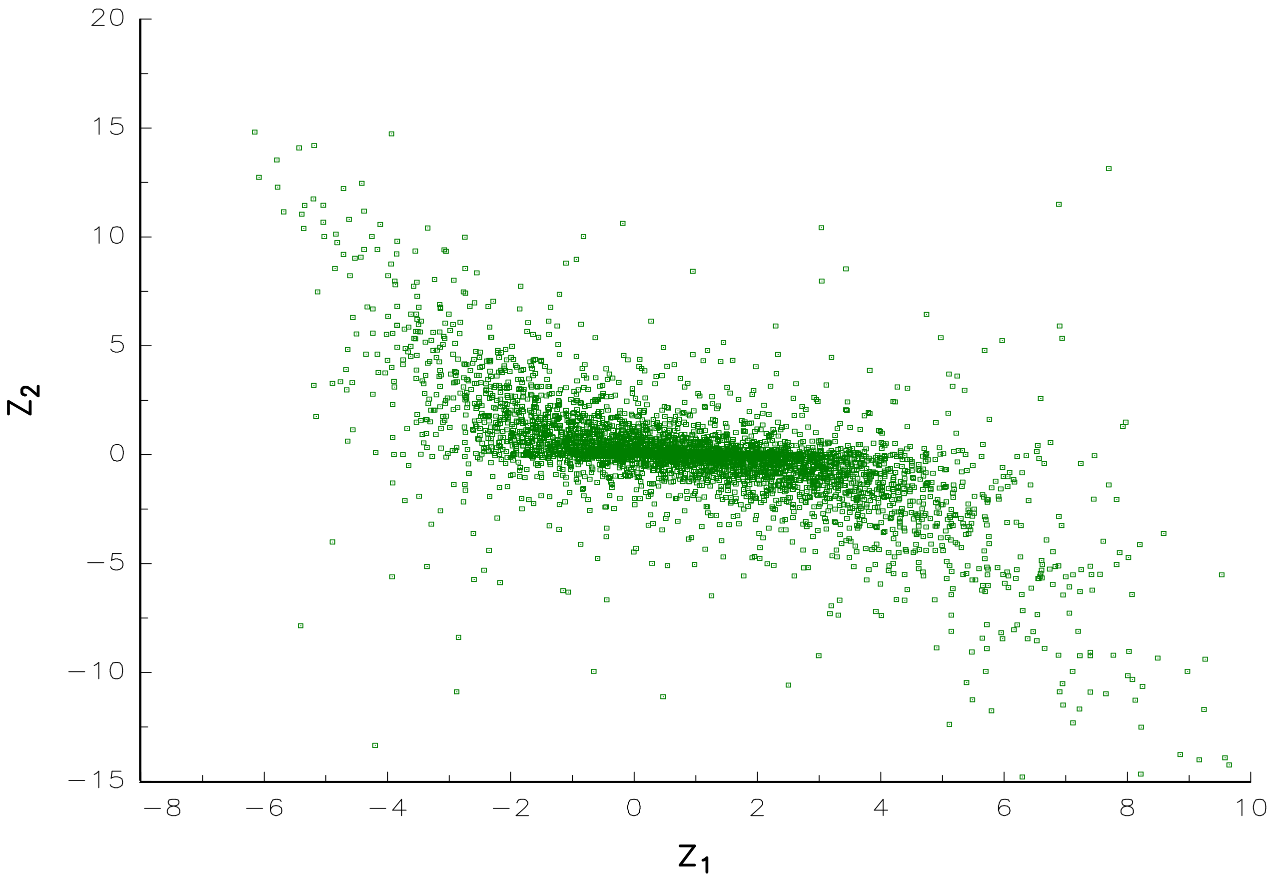


Figure: Simulation of the random vector (Z_1, Z_2)

Random matrices

- Orthogonal and covariance matrices
- Correlation matrices
- Wishart matrices

⇒ HFRM, Chapter 13, Section 13.1.4, pages 807-813

Brownian motion

- A Brownian motion (or a Wiener process) is a stochastic process $W(t)$, whose increments are stationary and independent:

$$W(t) - W(s) \sim \mathcal{N}(0, t - s)$$

- We have:

$$\begin{cases} W(0) = 0 \\ W(t) = W(s) + \epsilon(s, t) \end{cases}$$

  where $\epsilon(s, t) \sim \mathcal{N}(0, t - s)$ are *iid* random variables
- To simulate $W(t)$ at different dates $t_1, t_2, \ldots$, we have:

$$W_{m+1} = W_m + \sqrt{t_{m+1} - t_m} \cdot \varepsilon_m$$

  where W_m is the numerical realization of $W(t_m)$ and $\varepsilon_m \sim \mathcal{N}(0, 1)$ are *iid* random variables
- In the case of fixed-interval times $t_{m+1} - t_m = h$, we obtain the recursion:

$$W_{m+1} = W_m + \sqrt{h} \cdot \varepsilon_m$$

Geometric Brownian motion

- The geometric Brownian motion is described by the following SDE:

$$\left\{ \begin{array}{l} \mathrm{d}X(t) = \mu X(t)\,\mathrm{d}t + \sigma X(t)\,\mathrm{d}W(t) \\ X(0) = x_0 \end{array} \right.$$

- Its solution is given by:

$$X(t) = x_0 \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right) = g(W(t))$$

Geometric Brownian motion

1. Simulating the geometric Brownian motion $X(t)$ can be done by applying the transform method to the process $W(t)$
2. Another approach to simulate $X(t)$ consists in using the following formula:

$$X(t) = X(s) \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma\left(W(t) - W(s)\right)\right)$$

   We have:

$$X_{m+1} = X_m \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t_{m+1} - t_m) + \sigma\sqrt{t_{m+1} - t_m} \cdot \varepsilon_m\right)$$

   where $X_m = X(t_m)$ and $\varepsilon_m \sim \mathcal{N}(0,1)$ are *iid* random variables
3. If we consider fixed-interval times, the numerical realization becomes:

$$X_{m+1} = X_m \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)h + \sigma\sqrt{h} \cdot \varepsilon_m\right)$$

Geometric Brownian motion

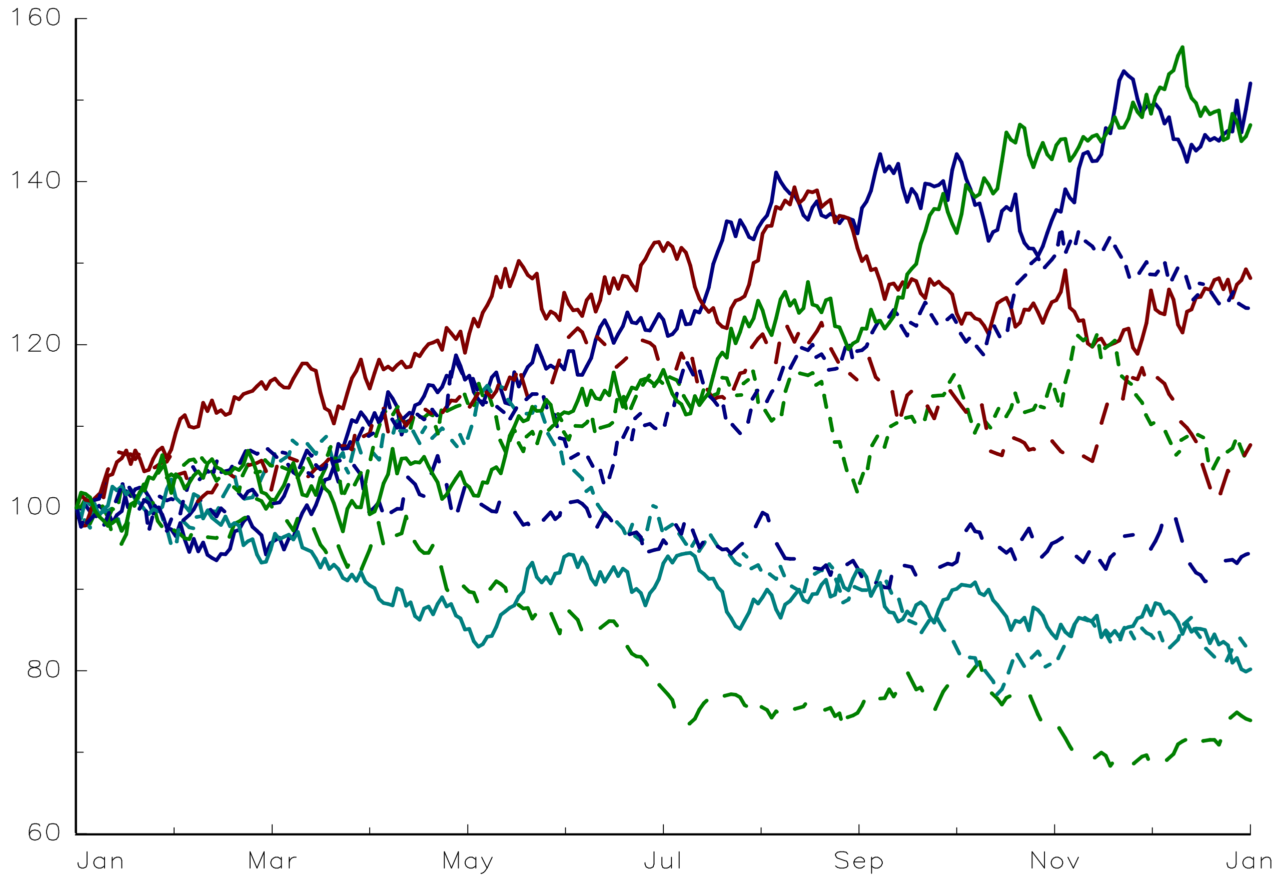


Figure: Simulation of the geometric Brownian motion

Ornstein-Uhlenbeck process

- The stochastic differential equation of the Ornstein-Uhlenbeck process is:

$$\begin{cases} \mathrm{d}X(t) = a(b - X(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W(t) \\ X(0) = x_0 \end{cases}$$

- The solution of the SDE is:

$$X(t) = x_0 e^{-at} + b\left(1 - e^{-at}\right) + \sigma \int_0^t e^{a(\theta - t)}\,\mathrm{d}W(\theta)$$

- We also have:

$$X(t) = X(s)\,e^{-a(t-s)} + b\left(1 - e^{-a(t-s)}\right) + \sigma \int_s^t e^{a(\theta - t)}\,\mathrm{d}W(\theta)$$

where:

$$\int_s^t e^{a(\theta - t)}\,\mathrm{d}W(\theta) \sim \mathcal{N}\left(0, \frac{1 - e^{-2a(t-s)}}{2a}\right)$$

Ornstein-Uhlenbeck process

If we consider fixed-interval times, we obtain the following simulation scheme:

$$X_{m+1} = X_m e^{-ah} + b\left(1 - e^{-ah}\right) + \sigma \sqrt{\frac{1 - e^{-2ah}}{2a}} \cdot \varepsilon_m$$

where $\varepsilon_m \sim \mathcal{N}(0,1)$ are *iid* random variables

Ornstein-Uhlenbeck process

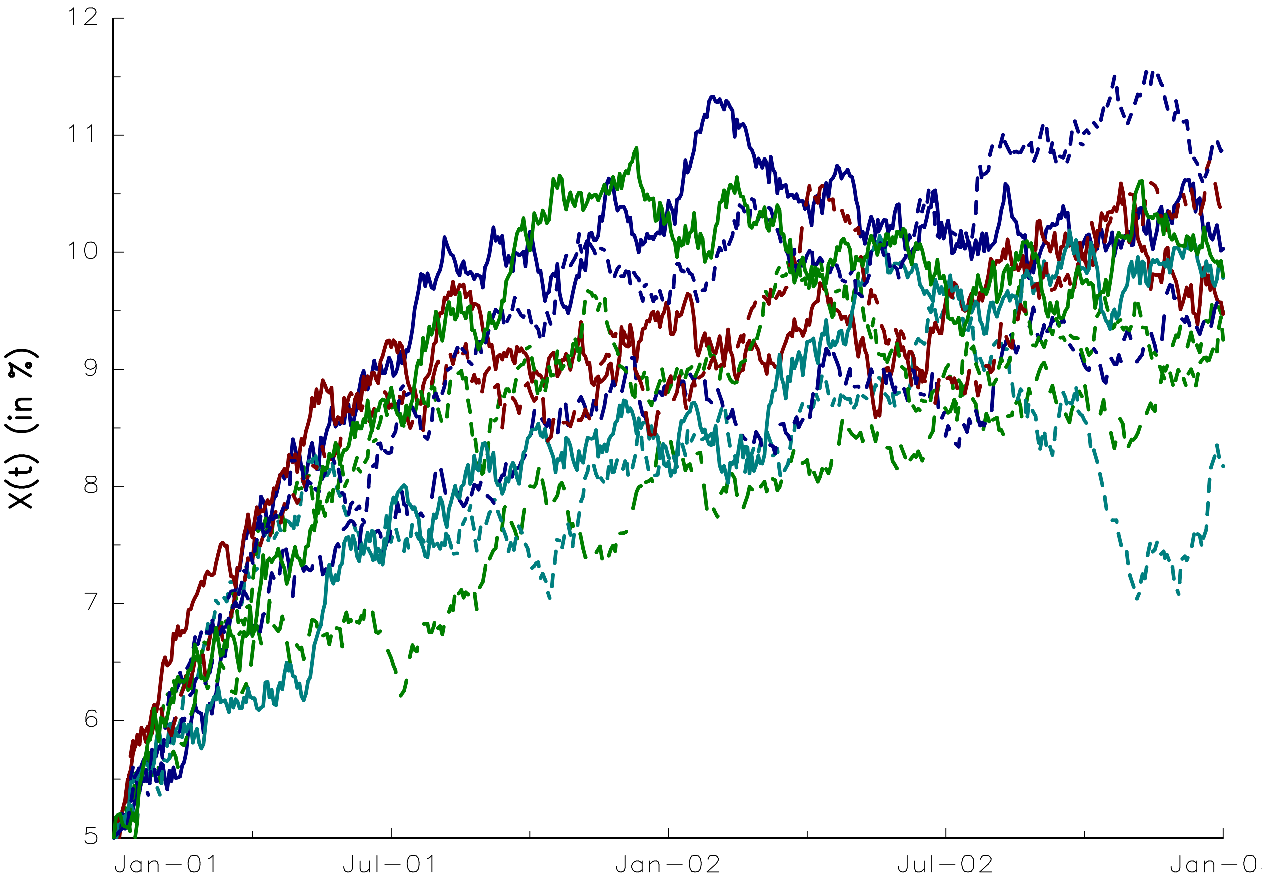


Figure: Simulation of the Ornstein-Uhlenbeck process

Stochastic differential equations without an explicit solution

- Let $X(t)$ be the solution of the following SDE:

$$\left\{\begin{array}{l} \mathrm{d}X(t) = \mu(t, X)\,\mathrm{d}t + \sigma(t, X)\,\mathrm{d}W(t) \\ X(0) = x_0 \end{array}\right.$$

- The Euler-Maruyama scheme uses the following approximation:

$$X(t) - X(s) \approx \mu(t, X(s)) \cdot (t - s) + \sigma(t, X(s)) \cdot (W(t) - W(s))$$

- If we consider fixed-interval times, the Euler-Maruyama scheme becomes:

$$X_{m+1} = X_m + \mu(t_m, X_m)\,h + \sigma(t_m, X_m)\sqrt{h} \cdot \varepsilon_m$$

where $\varepsilon_m \sim \mathcal{N}(0, 1)$ are *iid* random variables

Stochastic differential equations without an explicit solution

The fixed-interval Milstein scheme is:

$$\begin{aligned} X_{m+1} &= X_m + \mu\left(t_m, X_m\right) h + \sigma\left(t_m, X_m\right)\sqrt{h}\cdot\varepsilon_m + \\ & \frac{1}{2}\sigma\left(t_m, X_m\right)\partial_x\sigma\left(t_m, X_m\right) h\left(\varepsilon_m^2 - 1\right) \end{aligned}$$

Stochastic differential equations without an explicit solution

If we consider the geometric Brownian motion, the Euler-Maruyama scheme is:

$$X_{m+1} = X_m + \mu X_m h + \sigma X_m \sqrt{h} \cdot \varepsilon_m$$

whereas the Milstein scheme is:

$$\begin{aligned} X_{m+1} &= X_m + \mu X_m h + \sigma X_m \sqrt{h} \cdot \varepsilon_m + \frac{1}{2}\sigma^2 X_m h \left(\varepsilon_m^2 - 1\right) \\ &= X_m + \left(\mu - \frac{1}{2}\sigma^2\right) X_m h + \sigma X_m \sqrt{h} \left(1 + \frac{1}{2}\sigma\sqrt{h}\varepsilon_m\right)\varepsilon_m \end{aligned}$$

Stochastic differential equations without an explicit solution

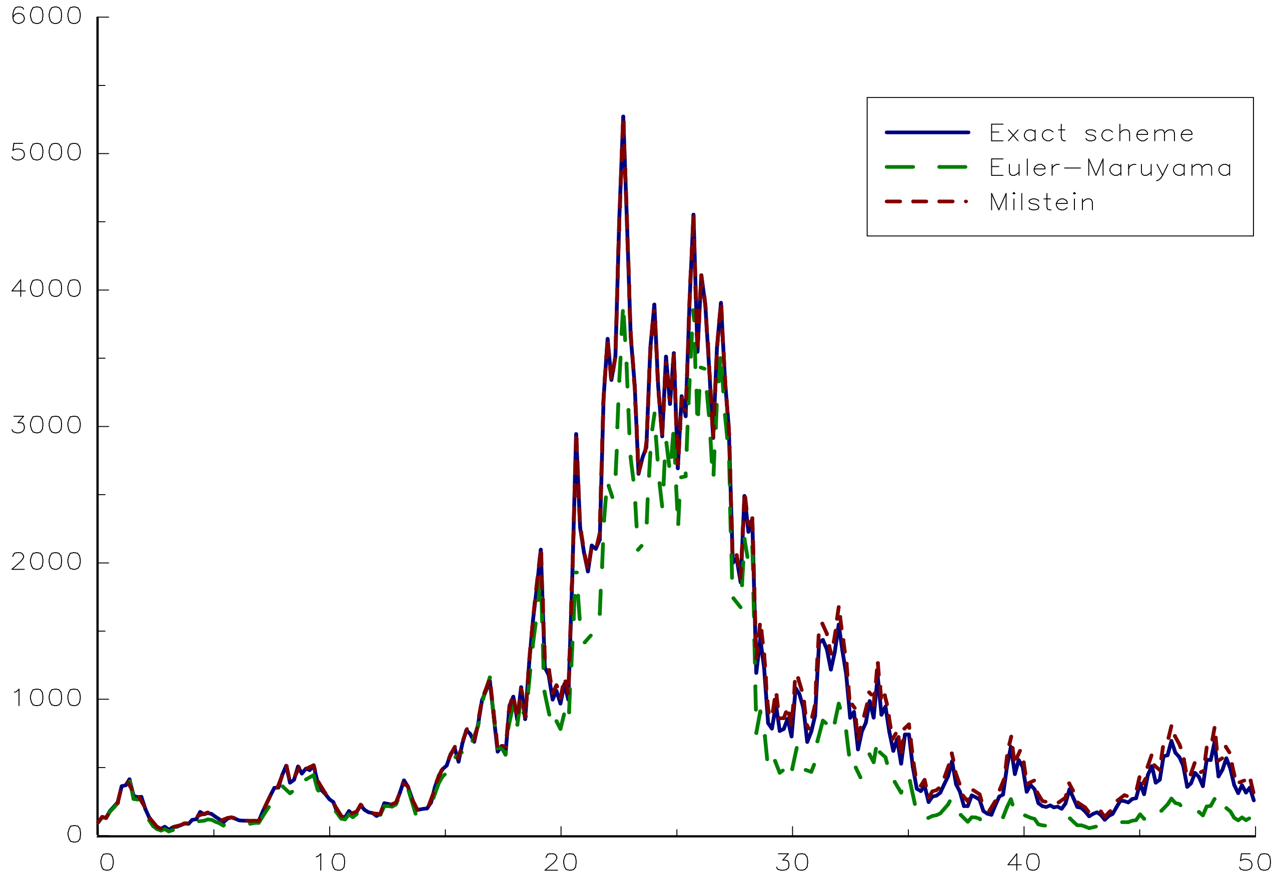


Figure: Comparison of exact, Euler-Maruyama and Milstein schemes (monthly discretization)

Stochastic differential equations without an explicit solution

When we don't know the analytical solution of $X(t)$, it is natural to simulate the numerical solution of $X(t)$ using Euler-Maruyama and Milstein schemes. However, it may be sometimes more efficient to find the numerical solution of $Y(t) = f(t, X(t))$ instead of $X(t)$ itself, in particular when $Y(t)$ is more regular than $X(t)$

Stochastic differential equations without an explicit solution

- By Itô's lemma, we have:

$$\begin{aligned} \mathrm{d}Y(t) &= \left(\partial_t f(t,X) + \mu(t,X)\,\partial_x f(t,X) + \frac{1}{2}\sigma^2(t,X)\,\partial_x^2 f(t,X)\right)\mathrm{d}t + \\ & \quad \sigma(t,X)\,\partial_x f(t,X)\,\mathrm{d}W(t) \end{aligned}$$

- By using the inverse function $X(t) = f^{-1}(t, Y(t))$, we obtain:

$$\mathrm{d}Y(t) = \mu'(t,Y)\,\mathrm{d}t + \sigma'(t,Y)\,\mathrm{d}W(t)$$

  where $\mu'(t,Y)$ and $\sigma'(t,Y)$ are functions of $\mu(t,X)$, $\sigma(t,X)$ and $f(t,X)$

- We can then simulate the solution of $Y(t)$ using an approximation scheme and deduce the numerical solution of $X(t)$ by applying the transformation method:

$$X_m = f^{-1}(t_m, Y_m)$$

Stochastic differential equations without an explicit solution

Let us consider the geometric Brownian motion $X(t)$. The solution of $Y(t) = \ln X(t)$ is equal to:

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

We deduce that the Euler-Maruyama (or Milstein) scheme with fixed-interval times is:

$$Y_{m+1} = Y_m + \left(\mu - \frac{1}{2}\sigma^2\right)h + \sigma\sqrt{h}\cdot\varepsilon_m$$

It follows that:

$$\ln X_{m+1} = \ln X_m + \left(\mu - \frac{1}{2}\sigma^2\right)h + \sigma\sqrt{h}\cdot\varepsilon_m$$

Stochastic differential equations without an explicit solution

The CIR process is $\mathrm{d}X(t) = (\alpha + \beta X(t))\,\mathrm{d}t + \sigma\sqrt{X(t)}\,\mathrm{d}W(t)$. Using the transformation $Y(t) = \sqrt{X(t)}$, we obtain the following SDE:

$$\begin{aligned}
\mathrm{d}Y(t) &= \left(\frac{1}{2}\frac{(\alpha+\beta X(t))}{\sqrt{X(t)}} - \frac{1}{8}\frac{\sigma^2 X(t)}{X(t)^{3/2}}\right)\mathrm{d}t + \frac{1}{2}\frac{\sigma\sqrt{X(t)}}{\sqrt{X(t)}}\,\mathrm{d}W(t) \\
&= \frac{1}{2Y(t)}\left(\alpha + \beta Y^2(t) - \frac{1}{4}\sigma^2\right)\mathrm{d}t + \frac{1}{2}\sigma\,\mathrm{d}W(t)
\end{aligned}$$

We deduce that the Euler-Maruyama scheme of $Y(t)$ is:

$$Y_{m+1} = Y_m + \frac{1}{2Y_m}\left(\alpha + \beta Y_m^2 - \frac{1}{4}\sigma^2\right)h + \frac{1}{2}\sigma\sqrt{h}\cdot\varepsilon_m$$

It follows that:

$$X_{m+1} = \left(\sqrt{X_m} + \frac{1}{2\sqrt{X_m}}\left(\alpha + \beta X_m - \frac{1}{4}\sigma^2\right)h + \frac{1}{2}\sigma\sqrt{h}\cdot\varepsilon_m\right)^2$$

Poisson process

Let t_m be the time when the m^{th} event occurs. The numerical algorithm is then:

1. we set $t_0 = 0$ and $N(t_0) = 0$
2. we generate a uniform random variate u and calculate the random variate $e \sim \mathcal{E}(\lambda)$ with the formula:
$$e = -\frac{\ln u}{\lambda}$$
3. we update the Poisson process with:
$$t_{m+1} \leftarrow t_m + e \qquad \text{and} \qquad N(t_{m+1}) \leftarrow N(t_m) + 1$$
4. we go back to step 2

Mixed Poisson process (MPP)

The algorithm is initialized with a realization λ of the random intensity Λ

Non-homogenous Poisson process (NHPP)

- $\lambda(t)$ varies with time
- The inter-arrival times remain independent and exponentially distributed with:

$$\Pr\{T_1 > t\} = \exp(-\Lambda(t))$$

where T_1 is the duration of the first event and $\Lambda(t)$ is the integrated intensity function:

$$\Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s$$

- It follows that:

$$\Pr\left\{T_1 > \Lambda^{-1}(t)\right\} = \exp(-t) \Leftrightarrow \Pr\{\Lambda(T_1) > t\} = \exp(-t)$$

Non-homogenous Poisson process (NHPP)

We deduce that if $\{t_1, t_2, \ldots, t_M\}$ are the occurrence times of the NHPP of intensity $\lambda(t)$, then $\{\Lambda(t_1), \Lambda(t_2), \ldots, \Lambda(t_M)\}$ are the occurrence times of the homogeneous Poisson process (HPP) of intensity one. Therefore, the algorithm is:

1. we simulate t'_m the time arrivals of the homogeneous Poisson process with intensity $\lambda = 1$
2. we apply the transform $t_m = \Lambda^{-1}(t'_m)$

Non-homogenous Poisson process (NHPP)

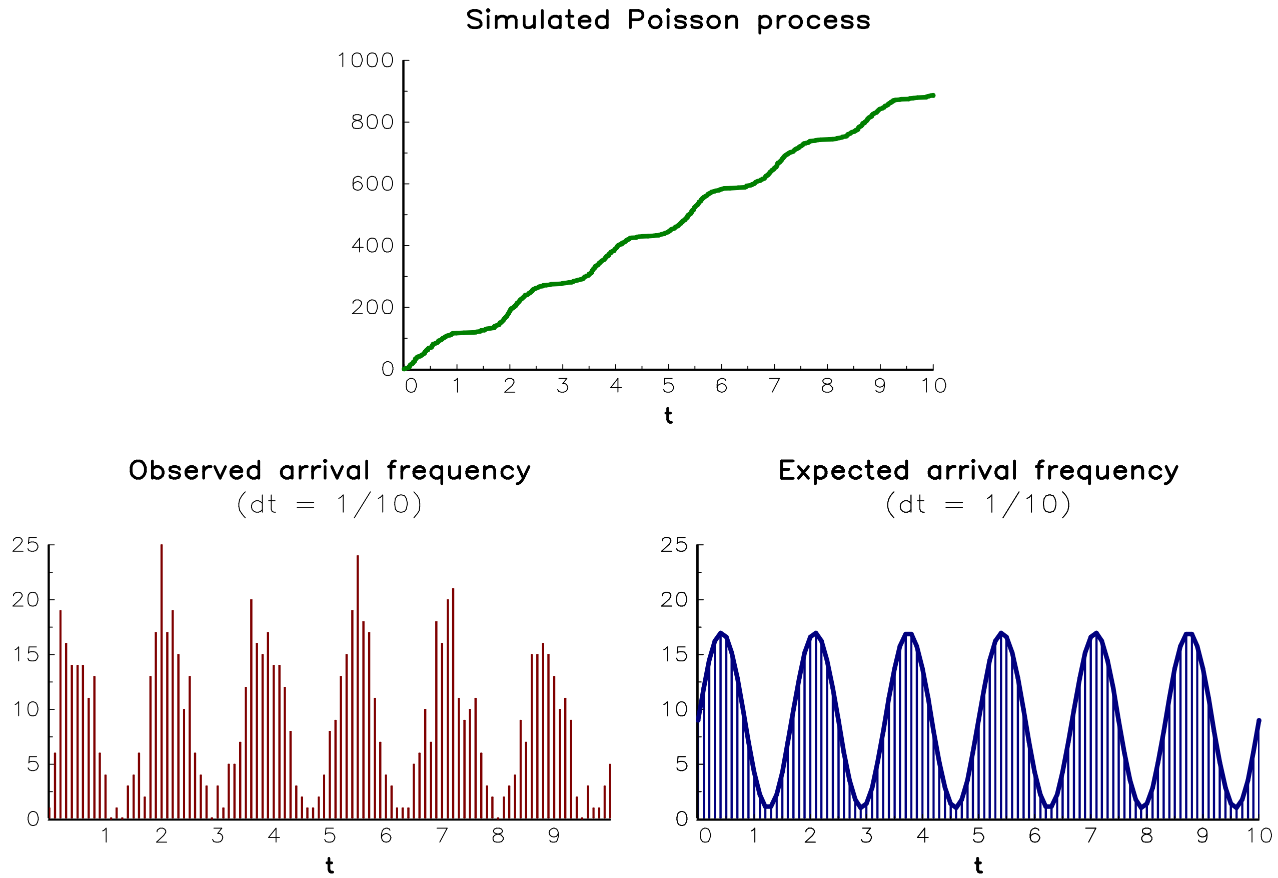


Figure: Simulation of a non-homogenous Poisson process with cyclical intensity

Multidimensional Brownian motion

- Let $W(t) = (W_1(t), \ldots, W_n(t))$, be a n-dimensional Brownian motion
- Each component $W_i(t)$ is a Brownian motion:

$$W_i(t) - W_i(s) \sim \mathcal{N}(0, t - s)$$

- We have:

$$\mathbb{E}[W_i(t) W_j(s)] = \min(t, s) \cdot \rho_{i,j}$$

where $\rho_{i,j}$ is the correlation between the two Brownian motions W_i and W_j

- We deduce that:

$$\left\{ \begin{array}{l} W(0) = \mathbf{0} \\ W(t) = W(s) + \epsilon(s, t) \end{array} \right.$$

where $\epsilon(s, t) \sim \mathcal{N}_n(\mathbf{0}, (t - s)\ \rho)$ are *iid* random vectors

Multidimensional Brownian motion

- It follows that the numerical solution is:

$$W_{m+1} = W_m + \sqrt{t_{m+1} - t_m} \cdot P \cdot \varepsilon_m$$

  where P is the Cholesky decomposition of the correlation matrix ρ and $\varepsilon_m \sim \mathcal{N}_n(0, I)$ are *iid* random vectors

- In the case of fixed-interval times, the recursion becomes:

$$W_{m+1} = W_m + \sqrt{h} \cdot P \cdot \varepsilon_m$$

Multidimensional Brownian motion

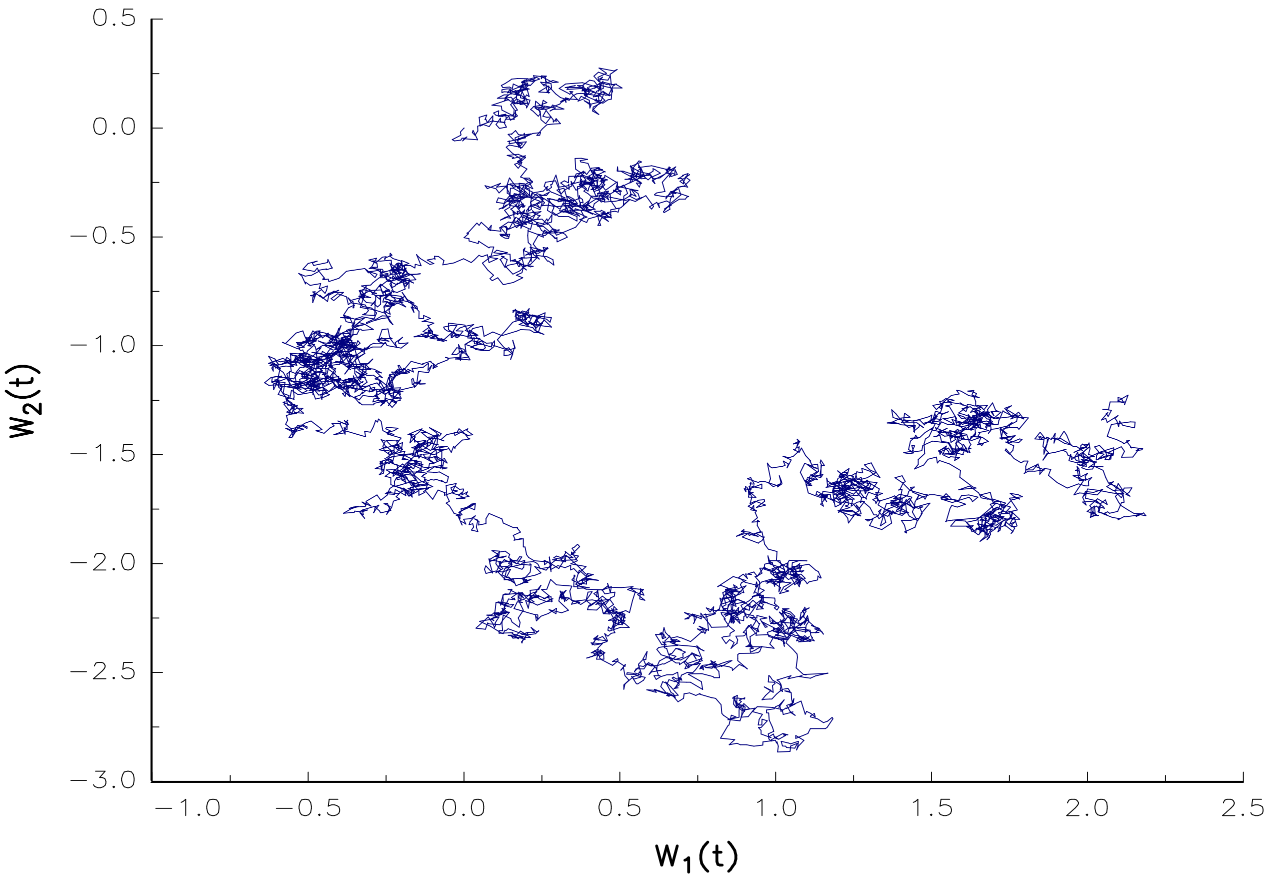


Figure: Brownian motion in the plane (independent case)

Multidimensional Brownian motion

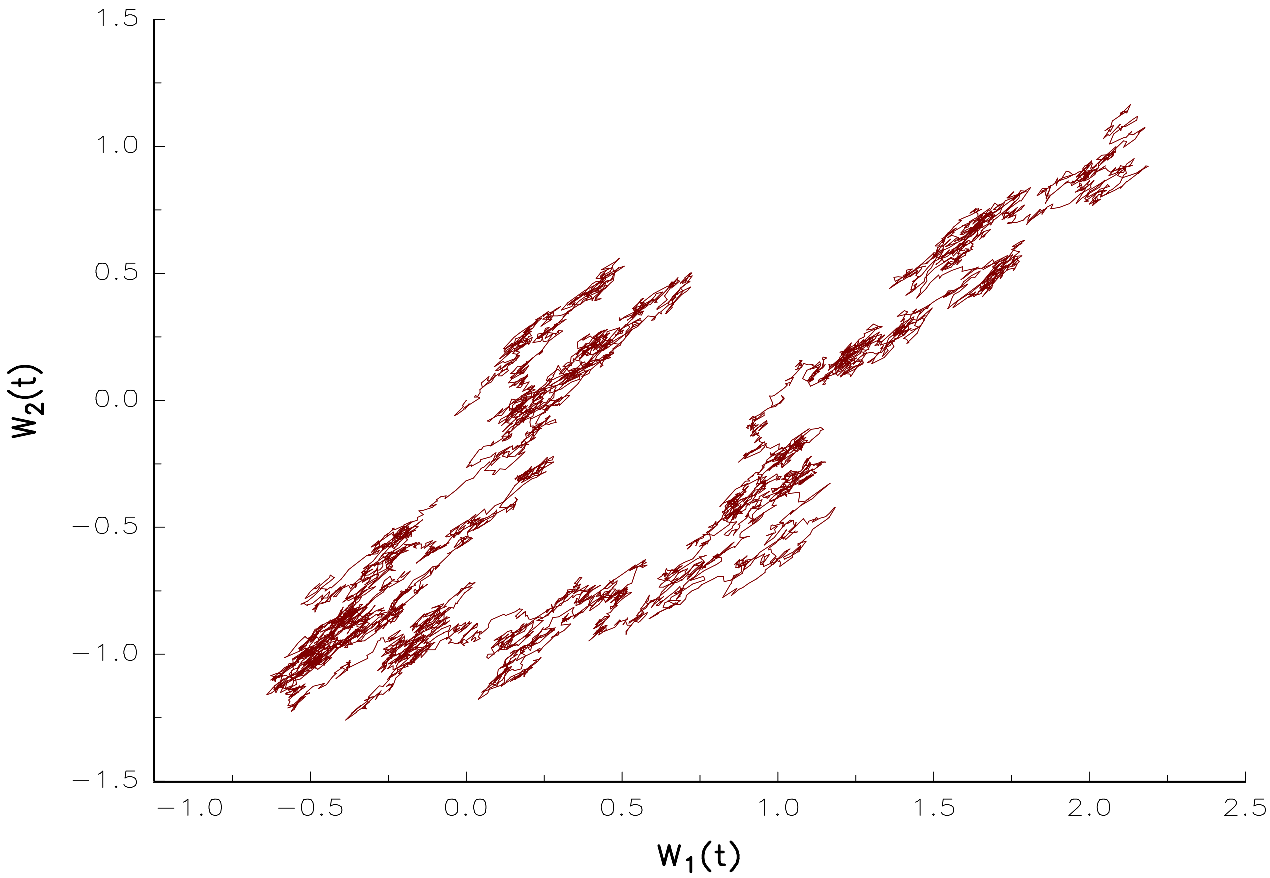


Figure: Brownian motion in the plane ($\rho_{1,2} = 85\%$)

Multidimensional geometric Brownian motion

- We consider the multidimensional geometric Brownian motion:

$$\begin{cases} \mathrm{d}X(t) = \mu \odot X(t)\,\mathrm{d}t + \operatorname{diag}(\sigma \odot X(t))\,\mathrm{d}W(t) \\ X(0) = x_0 \end{cases}$$

  where $X(t) = (X_1(t), \ldots, X_n(t))$, $\mu = (\mu_1, \ldots, \mu_n)$, $\sigma = (\sigma_1, \ldots, \sigma_n)$ and $W(t) = (W_1(t), \ldots, W_n(t))$ is a n-dimensional Brownian motion with $\mathbb{E}\left[W(t)\,W(t)^\top\right] = \rho\, t$

- If we consider the j^{th} component of $X(t)$, we have:

$$\mathrm{d}X_j(t) = \mu_j X_j(t)\,\mathrm{d}t + \sigma_j X_j(t)\,\mathrm{d}W_j(t)$$

- The solution of the multidimensional SDE is a multivariate log-normal process with:

$$X_j(t) = X_j(0) \cdot \exp\left(\left(\mu_j - \frac{1}{2}\sigma_j^2\right)t + \sigma_j W_j(t)\right)$$

  where $W(t) \sim \mathcal{N}_n(0, \rho\, t)$

Multidimensional geometric Brownian motion

- We deduce that the exact scheme to simulate the multivariate GBM is:

$$\left\{\begin{array}{l} X_{1,m+1} = X_{1,m} \cdot \exp\left(\left(\mu_1 - \frac{1}{2}\sigma_1^2\right)(t_{m+1} - t_m) + \sigma_1\sqrt{t_{m+1} - t_m} \cdot \varepsilon_{1,m}\right) \\ \vdots \\ X_{j,m+1} = X_{j,m} \cdot \exp\left(\left(\mu_j - \frac{1}{2}\sigma_j^2\right)(t_{m+1} - t_m) + \sigma_j\sqrt{t_{m+1} - t_m} \cdot \varepsilon_{j,m}\right) \\ \vdots \\ X_{n,m+1} = X_{n,m} \cdot \exp\left(\left(\mu_n - \frac{1}{2}\sigma_n^2\right)(t_{m+1} - t_m) + \sigma_n\sqrt{t_{m+1} - t_m} \cdot \varepsilon_{n,m}\right) \end{array}\right.$$

where $(\varepsilon_{1,m}, \ldots, \varepsilon_{n,m}) \sim \mathcal{N}_n(\mathbf{0}, \rho)$

Euler-Maruyama and Milstein schemes

- We consider the general SDE:

$$\left\{\begin{array}{l} \mathrm{d}X(t) = \mu(t, X(t))\,\mathrm{d}t + \sigma(t, X(t))\,\mathrm{d}W(t) \\ X(0) = x_0 \end{array}\right.$$

  where $X(t)$ and $\mu(t, X(t))$ are $n \times 1$ vectors, $\sigma(t, X(t))$ is a $n \times p$ matrix and $W(t)$ is a $p \times 1$ vector

- We assume that $\mathbb{E}\left[W(t)\,W(t)^{\top}\right] = \rho\,t$, where ρ is a $p \times p$ correlation matrix

Euler-Maruyama and Milstein schemes

- The corresponding Euler-Maruyama scheme is:

$$X_{m+1} = X_m + \mu\left(t_m, X_m\right) \cdot \left(t_{m+1} - t_m\right) + \sigma\left(t_m, X_m\right)\sqrt{t_{m+1} - t_m} \cdot \varepsilon_m$$

where $\varepsilon_m \sim \mathcal{N}_p\left(0, \rho\right)$

- In the case of a diagonal system, we retrieve the one-dimensional scheme:

$$X_{j,m+1} = X_{j,m} + \mu_j\left(t_m, X_{j,m}\right) \cdot \left(t_{m+1} - t_m\right) + \sigma_{j,j}\left(t_m, X_{j,m}\right) \cdot \sqrt{t_{m+1} - t_m}\varepsilon_{j,m}$$

However, the random variables $\varepsilon_{j,m}$ and $\varepsilon_{j',m}$ may be correlated

Euler-Maruyama and Milstein schemes

We consider the Heston model:

$$\left\{\begin{array}{l} \mathrm{d}X(t) = \mu X(t)\,\mathrm{d}t + \sqrt{v(t)}X(t)\,\mathrm{d}W_1(t) \\ \mathrm{d}v(t) = a(b - v(t))\,\mathrm{d}t + \sigma\sqrt{v(t)}\,\mathrm{d}W_2(t) \end{array}\right.$$

where $\mathbb{E}[W_1(t)\,W_2(t)] = \rho\,t$. By applying the fixed-interval Euler-Maruyama scheme to $(\ln X(t), v(t))$, we obtain:

$$\ln X_{m+1} = \ln X_m + \left(\mu - \frac{1}{2}v_m\right)h + \sqrt{v_m h} \cdot \varepsilon_{1,m}$$

and:

$$v_{m+1} = v_m + a(b - v_m)\,h + \sigma\sqrt{v_m h} \cdot \varepsilon_{2,m}$$

Here, $\varepsilon_{1,m}$ and $\varepsilon_{2,m}$ are two standard Gaussian random variables with correlation ρ

Euler-Maruyama and Milstein schemes

The multidimensional version of the Milstein scheme is:

$$\begin{aligned} X_{j,m+1} &= X_{j,m} + \mu_j \left(t_m, X_m\right)\left(t_{m+1} - t_m\right) + \sum_{k=1}^{p} \sigma_{j,k}\left(t_m, X_m\right)\Delta W_{k,m} + \\ & \sum_{k=1}^{p}\sum_{k'=1}^{p} \mathcal{L}^{(k)}\sigma_{j,k'}\left(t_m, X_m\right)\mathcal{I}_{(k,k')} \end{aligned}$$

where $\Delta W_{k,m} = W_k\left(t_{m+1}\right) - W_k\left(t_m\right)$ and:

$$\mathcal{L}^{(k)} f\left(t, x\right) = \sum_{k''=1}^{n} \sigma_{k'',k}\left(t_m, X_m\right)\frac{\partial\, f\left(t, x\right)}{\partial\, x_{k''}}$$

and:

$$\mathcal{I}_{(k,k')} = \int_{t_m}^{t_{m+1}}\int_{t_m}^{s} \mathrm{d}W_k\left(t\right)\mathrm{d}W_{k'}\left(s\right)$$

Euler-Maruyama and Milstein schemes

In the case of a diagonal system, the Milstein scheme may be simplified as follows:

$$\begin{aligned} X_{j,m+1} &= X_{j,m} + \mu_j \left(t_m, X_{j,m}\right)\left(t_{m+1} - t_m\right) + \sigma_{j,j}\left(t_m, X_{j,m}\right)\Delta W_{j,m} + \\ & \quad \mathcal{L}^{(j)}\sigma_{j,j}\left(t_m, X_{j,m}\right)\mathcal{I}_{(j,j)} \end{aligned}$$

where:

$$\begin{aligned} \mathcal{I}_{(j,j)} &= \int_{t_m}^{t_{m+1}} \int_{t_m}^{s} \mathrm{d}W_j\left(t\right)\mathrm{d}W_j\left(s\right) \\ &= \int_{t_m}^{t_{m+1}} \left(W_j\left(s\right) - W_j\left(t_m\right)\right)\mathrm{d}W_j\left(s\right) \\ &= \frac{1}{2}\left(\left(\Delta W_{j,m}\right)^2 - \left(t_{m+1} - t_m\right)\right) \end{aligned}$$

Euler-Maruyama and Milstein schemes

We deduce that the Milstein scheme is:

$$\begin{aligned} X_{j,m+1} &= X_{j,m} + \mu_j \left(t_m, X_{j,m}\right) \left(t_{m+1} - t_m\right) + \\ & \sigma_{j,j} \left(t_m, X_{j,m}\right) \sqrt{t_{m+1} - t_m} \varepsilon_{j,m} + \\ & \frac{1}{2} \sigma_{j,j} \left(t_m, X_{j,m}\right) \partial_{x_j} \sigma_{j,j} \left(t_m, X_{j,m}\right) \left(t_{m+1} - t_m\right) \left(\varepsilon_{j,m}^2 - 1\right) \end{aligned}$$

Euler-Maruyama and Milstein schemes

If we apply the fixed-interval Milstein scheme to the Heston model, we obtain:

$$\ln X_{m+1} = \ln X_m + \left(\mu - \frac{1}{2} v_m \right) h + \sqrt{v_m h} \cdot \varepsilon_{1,m}$$

and:

$$v_{m+1} = v_m + a \left(b - v_m \right) h + \sigma \sqrt{v_m h} \cdot \varepsilon_{2,m} + \frac{1}{4} \sigma^2 h \left(\varepsilon_{2,m}^2 - 1 \right)$$

Here, $\varepsilon_{1,m}$ and $\varepsilon_{2,m}$ are two standard Gaussian random variables with correlation ρ

Euler-Maruyama and Milstein schemes

Remark

The multidimensional Milstein scheme is generally not used, because the terms $\mathcal{L}^{(k)}\sigma_{j,k'}(t_m, X_m)\mathcal{I}_{(k,k')}$ are complicated to simulate. For the Heston model, we obtain a very simple scheme, because we only apply the Milstein scheme to the process $v(t)$ and not to the vector process $(\ln X(t), v(t))$

Euler-Maruyama and Milstein schemes

If we also apply the Milstein scheme to $\ln X(t)$, we obtain:

$$\ln X_{m+1} = \ln X_m + \left(\mu - \frac{1}{2} v_m\right) h + \sqrt{v_m h} \cdot \varepsilon_{1,m} + A_m$$

where:

$$\begin{aligned} A_m &= \sum_{k=1}^{2} \sum_{k'=1}^{2} \left(\sum_{k''=1}^{2} \sigma_{k'',k}(t_m, X_m) \frac{\sigma_{1,k'}(t_m, X_m)}{\partial x_{k''}} \right) \mathcal{I}_{(k,k')} \\ &= \sigma \sqrt{v(t)} \cdot \frac{1}{2\sqrt{v(t)}} \cdot \mathcal{I}_{(2,1)} \\ &= \frac{\sigma}{2} \cdot \mathcal{I}_{(2,1)} \end{aligned}$$

Euler-Maruyama and Milstein schemes

Let $W_2(t) = \rho W_1(t) + \sqrt{1-\rho^2} W^\star(t)$ where $W^\star(t)$ is a Brownian motion independent from $W_1(t)$. It follows that:

$$\begin{aligned}
\mathcal{I}_{(2,1)} &= \int_{t_m}^{t_{m+1}} \int_{t_m}^{s} \mathrm{d}W_2(t)\,\mathrm{d}W_1(s) \\
&= \int_{t_m}^{t_{m+1}} \left(\rho W_1(s) + \sqrt{1-\rho^2} W^\star(s)\right) \mathrm{d}W_1(s) - \\
&\quad \int_{t_m}^{t_{m+1}} \left(\rho W_1(t_m) + \sqrt{1-\rho^2} W^\star(t_m)\right) \mathrm{d}W_1(s) \\
&= \rho \int_{t_m}^{t_{m+1}} \left(W_1(s) - W_1(t_m)\right) \mathrm{d}W_1(s) + \\
&\quad \sqrt{1-\rho^2} \int_{t_m}^{t_{m+1}} \left(W^\star(s) - W^\star(t_m)\right) \mathrm{d}W_1(s)
\end{aligned}$$

and:

$$\mathcal{I}_{(2,1)} = \frac{1}{2}\rho\left(\left(\Delta W_{1,m}\right)^2 - \left(t_{m+1} - t_m\right)\right) + B_m$$

Euler-Maruyama and Milstein schemes

We finally deduce that the multidimensional Milstein scheme of the Heston model is:

$$\ln X_{m+1} = \ln X_m + \left(\mu - \frac{1}{2} v_m\right) h + \sqrt{v_m h} \cdot \varepsilon_{1,m} + \frac{1}{4} \rho \sigma h \left(\varepsilon_{1,m}^2 - 1\right) + B_m$$

and:

$$v_{m+1} = v_m + a \left(b - v_m\right) h + \sigma \sqrt{v_m h} \cdot \varepsilon_{2,m} + \frac{1}{4} \sigma^2 h \left(\varepsilon_{2,m}^2 - 1\right)$$

where B_m is a correction term defined by:

$$B_m = \sqrt{1 - \rho^2} \int_{t_m}^{t_{m+1}} \left(W^\star(s) - W^\star(t_m)\right) \,\mathrm{d}W_1(s)$$

A basic example

- Suppose we have a circle with radius r and a $2r \times 2r$ square of the same center. Since the area of the circle is equal to πr^2, the numerical calculation of π is equivalent to compute the area of the circle with $r = 1$
- In this case, the area of the square is 4, and we have:

$$\pi = 4 \frac{\mathcal{A}\left(\text{circle}\right)}{\mathcal{A}\left(\text{square}\right)}$$

- To determine π, we simulate n_S random vectors (u_s, v_s) of uniform random variables $\mathcal{U}_{[-1,1]}$ and we obtain:

$$\pi = \lim_{n_S \rightarrow \infty} 4 \frac{n_c}{n}$$

where n_c is the number of points (u_s, v_s) in the circle:

$$n_c = \sum_{s=1}^{n_S} \mathbb{1}\left\{u_s^2 + v_s^2 \leq r^2\right\}$$

A basic example

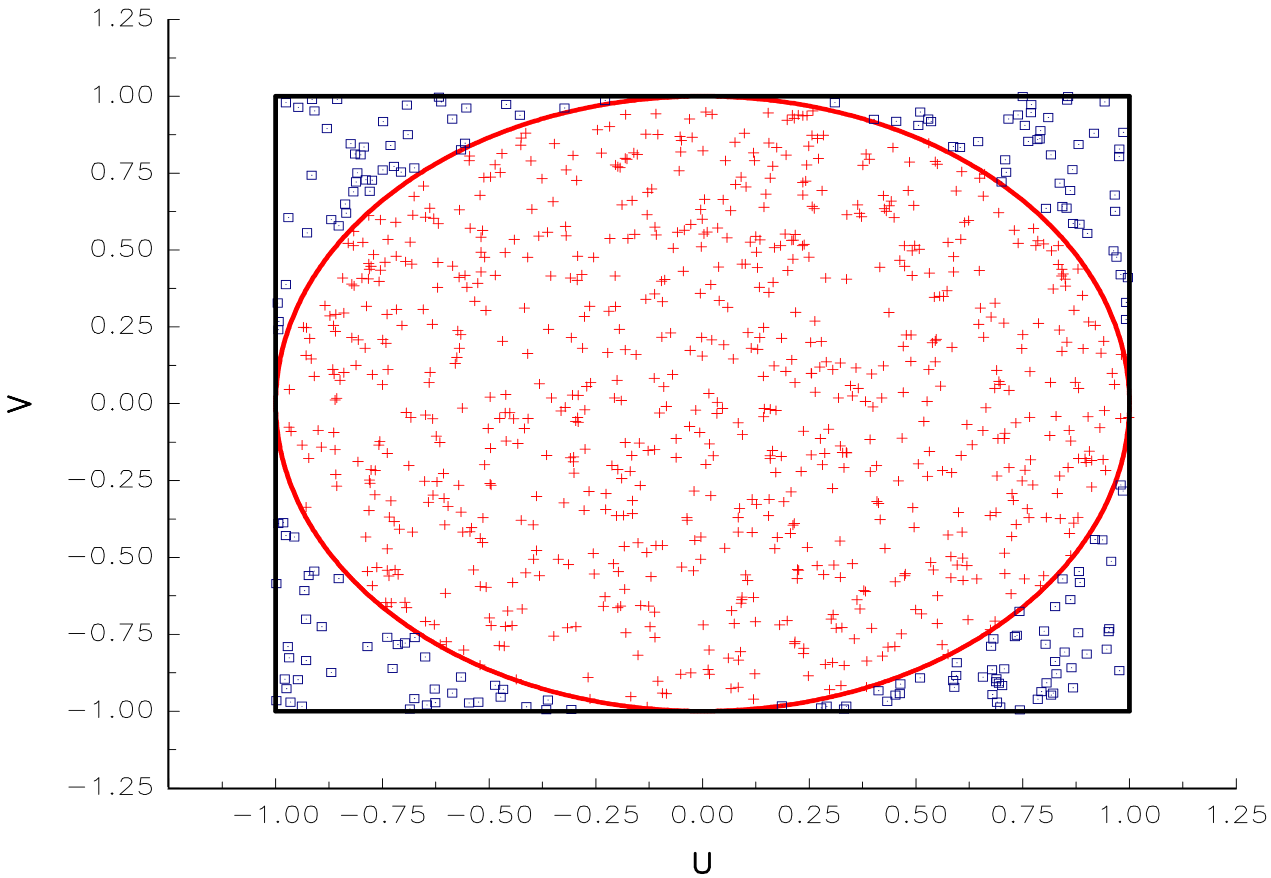


Figure: Computing π with 1 000 simulations

Theoretical framework

- We consider the multiple integral:

$$I = \int \cdots \int_{\Omega} \varphi\left(x_1, \ldots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

- Let $X = (X_1, \ldots, X_n)$ be a uniform random vector with probability distribution $\mathcal{U}_{[\Omega]}$, such that Ω is inscribed within the hypercube $[\Omega]$
- The pdf is:

$$f\left(x_1, \ldots, x_n\right) = 1$$

- We deduce that:

$$\begin{aligned} I &= \int \cdots \int_{[\Omega]} \mathbb{1}\left\{\left(x_1, \ldots, x_n\right) \in \Omega\right\} \cdot \varphi\left(x_1, \ldots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ &= \mathbb{E}\left[\mathbb{1}\left\{\left(X_1, \ldots, X_n\right) \in \Omega\right\} \cdot \varphi\left(X_1, \ldots, X_n\right)\right] \\ &= \mathbb{E}\left[h\left(X_1, \ldots, X_n\right)\right] \end{aligned}$$

where:

$$h\left(x_1, \ldots, x_n\right) = \mathbb{1}\left\{\left(x_1, \ldots, x_n\right) \in \Omega\right\} \cdot \varphi\left(x_1, \ldots, x_n\right)$$

Theoretical framework

- Let $\hat{I}_{n_S}$ be the random variable defined by:

$$\hat{I}_{n_S} = \frac{1}{n_S} \sum_{s=1}^{n_S} h\left(X_{1,s}, \ldots, X_{n,s}\right)$$

where $\left\{X_{1,s}, \ldots, X_{n,s}\right\}_{s \geq 1}$ is a sequence of *iid* random vectors with probability distribution $\mathcal{U}_{[\Omega]}$

- Using the strong law of large numbers, we obtain:

$$\begin{aligned} \lim_{n_s \to \infty} \hat{I}_{n_s} &= \mathbb{E}\left[h\left(X_1, \ldots, X_n\right)\right] \\ &= \int \cdots \int_{\Omega} \varphi\left(x_1, \ldots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \end{aligned}$$

- Moreover, the central limit theorem states that:

$$\lim_{n_s \to \infty} \sqrt{n_S} \left(\frac{\hat{I}_{n_s} - I}{\sigma\left(h\left(X_1, \ldots, X_n\right)\right)} \right) = \mathcal{N}\left(0, 1\right)$$

Theoretical framework

- When n_S is large, we can deduce the following confidence interval:

$$\left[\hat{I}_{n_S} - c_\alpha \cdot \frac{\hat{S}_{n_S}}{\sqrt{n_S}}, \hat{I}_{n_S} + c_\alpha \cdot \frac{\hat{S}_{n_S}}{\sqrt{n_S}}\right]$$

where α is the confidence level, $c_\alpha = \Phi^{-1}\left(\left(1+\alpha\right)/2\right)$ and $\hat{S}_{n_S}$ is the usual estimate of the standard deviation:

$$\hat{S}_{n_S} = \sqrt{\frac{1}{n_S - 1}\sum_{s=1}^{n_S} h^2\left(X_{1,s}, \ldots, X_{n,s}\right) - \hat{I}_{n_s}}$$

Theoretical framework

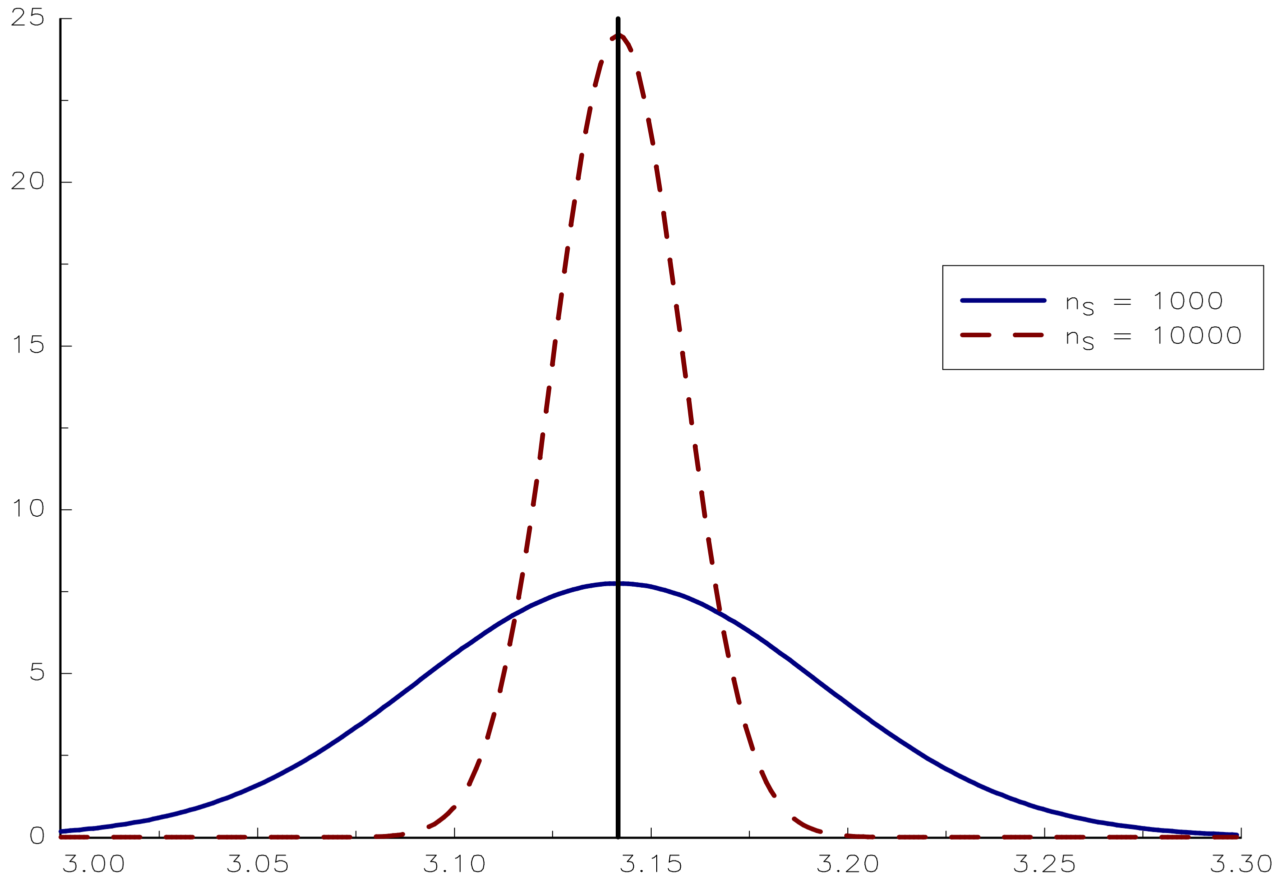


Figure: Density function of $\hat{\pi}_{n_S}$

Extension to the calculation of mathematical expectations

- Let $X = (X_1, \ldots, X_n)$ be a random vector with probability distribution $\mathbf{F}$. We have:

$$\begin{aligned}
\mathbb{E}\left[\varphi\left(X_1, \ldots, X_n\right)\right] &= \int \cdots \int \varphi\left(x_1, \ldots, x_n\right) \mathrm{d}\mathbf{F}\left(x_1, \cdots, x_n\right) \\
&= \int \cdots \int \varphi\left(x_1, \ldots, x_n\right) f\left(x_1, \cdots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \\
&= \int \cdots \int h\left(x_1, \ldots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n
\end{aligned}$$

  where f is the density function
- The Monte Carlo estimator of this integral is:

$$\hat{I}_{n_S} = \frac{1}{n_S} \sum_{s=1}^{n_S} \varphi\left(X_{1,s}, \ldots, X_{n,s}\right)$$

  where $\left\{X_{1,s}, \ldots, X_{n,s}\right\}_{s \geq 1}$ is a sequence of *iid* random vectors with probability distribution $\mathbf{F}$

Extension to the calculation of mathematical expectations

- The price of the look-back option with maturity T is given by:

$$\mathcal{C} = e^{-rT}\mathbb{E}\left[\left(S(T) - \min_{0\leq t\leq T} S(t)\right)^{+}\right]$$

- The price $S(t)$ of the underlying asset is given by the following SDE:

$$\mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}W(t)$$

where r is the interest rate and σ is the volatility of the asset

- For a given simulation s, we have:

$$S_{m+1}^{(s)} = S_m^{(s)} \cdot \exp\left(\left(r - \frac{1}{2}\sigma^2\right)(t_{m+1} - t_m) + \sigma\sqrt{t_{m+1} - t_m} \cdot \varepsilon_m^{(s)}\right)$$

where $\varepsilon_m^{(s)} \sim \mathcal{N}(0,1)$ and $T = t_M$

- The Monte Carlo estimator of the option price is then equal to:

$$\hat{\mathcal{C}} = \frac{e^{-rT}}{n_S}\sum_{s=1}^{n_S}\left(S_M^{(s)} - \min_m S_m^{(s)}\right)^{+}$$

Extension to the calculation of mathematical expectations

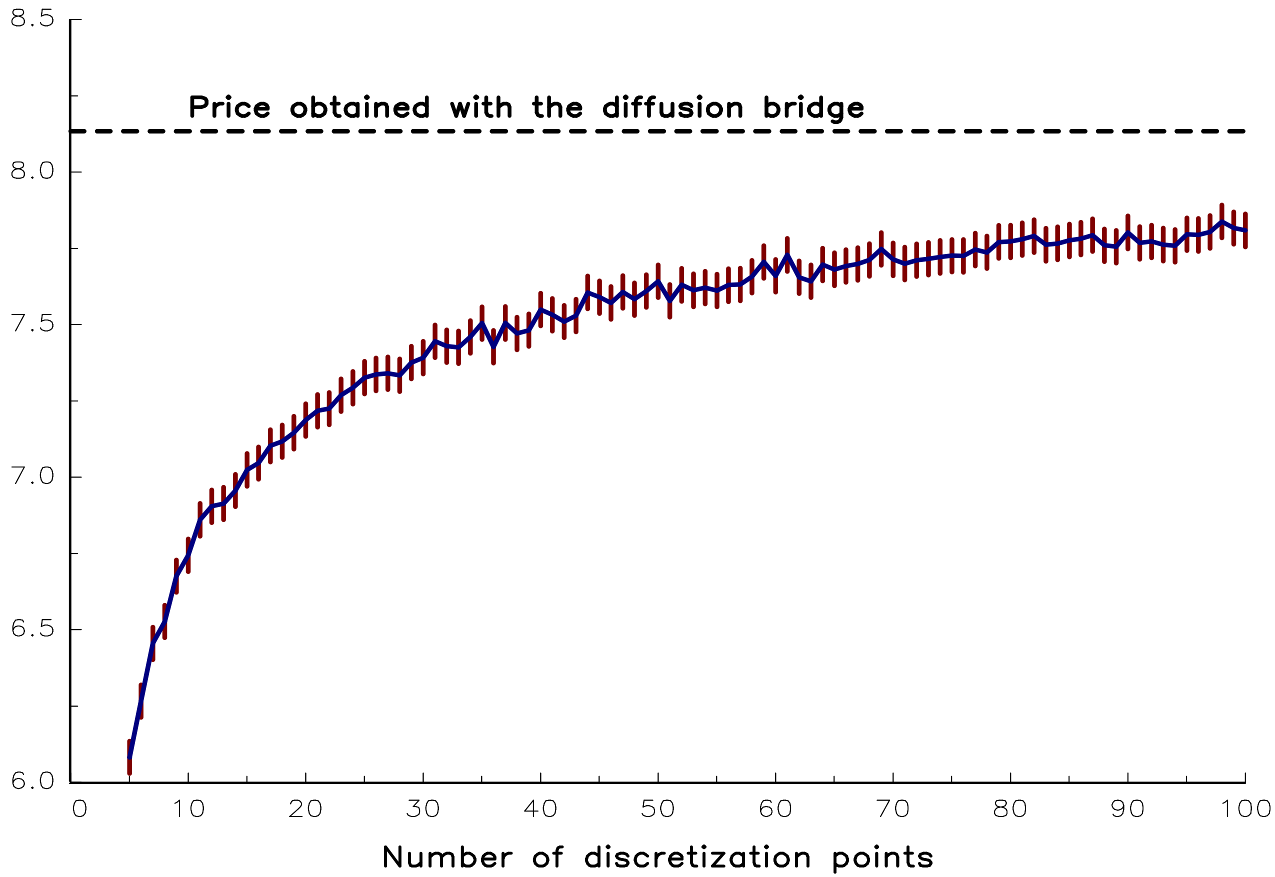


Figure: Computing the look-back option price

Extension to the calculation of mathematical expectations

- Let us consider the following integral:

$$I = \int \cdots \int h(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

- We can write it as follows:

$$I = \int \cdots \int \frac{h(x_1, \ldots, x_n)}{f(x_1, \cdots, x_n)} f(x_1, \cdots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

  where $f(x_1, \cdots, x_n)$ is a multidimensional density function

- We deduce that:

$$I = \mathbb{E}\left[\frac{h(X_1, \ldots, X_n)}{f(X_1, \ldots, X_n)}\right]$$

- This implies that we can compute an integral with the MC method by using any multidimensional distribution function

Extension to the calculation of mathematical expectations

If we apply this result to the calculation of π, we have:

$$\begin{aligned}\pi &= \iint_{x^2+y^2\leq 1} \mathrm{d}x\,\mathrm{d}y = \iint \mathbb{1}\left\{x^2+y^2\leq 1\right\}\mathrm{d}x\,\mathrm{d}y \\ &= \iint \frac{\mathbb{1}\left\{x^2+y^2\leq 1\right\}}{\phi(x)\,\phi(y)}\phi(x)\,\phi(y)\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

We deduce that:

$$\pi = \mathbb{E}\left[\frac{\mathbb{1}\left\{X^2+Y^2\leq 1\right\}}{\phi(X)\,\phi(Y)}\right]$$

where X and Y are two independent standard Gaussian random variables. We can then estimate π by:

$$\hat{\pi}_{n_S} = \frac{1}{n_S}\sum_{s=1}^{n_S}\frac{\mathbb{1}\left\{x_s^2+y_s^2\leq 1\right\}}{\phi(x_s)\,\phi(y_s)}$$

where x_s and y_s are two independent random variates from the probability distribution $\mathcal{N}(0,1)$

Extension to the calculation of mathematical expectations

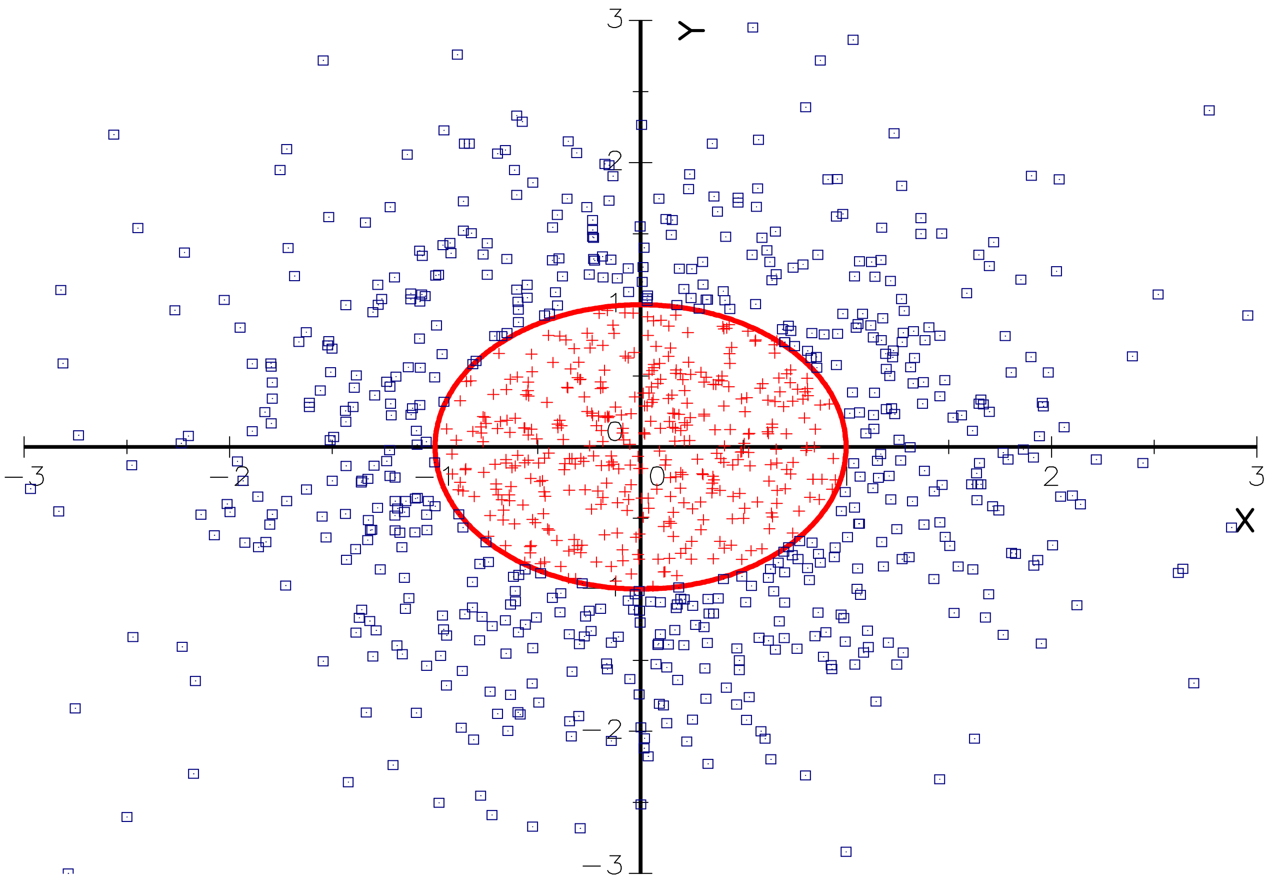


Figure: Computing *pi* with normal random numbers

Variance reduction

- We consider two unbiased estimators $\hat{I}_{n_S}^{(1)}$ and $\hat{I}_{n_S}^{(2)}$ of the integral I, meaning that $\mathbb{E}\left[\hat{I}_{n_S}^{(1)}\right] = \mathbb{E}\left[\hat{I}_{n_S}^{(2)}\right] = I$
- We say that $\hat{I}_{n_S}^{(1)}$ is more efficient than $\hat{I}_{n_S}^{(2)}$ if the inequality $\mathrm{var}\left(\hat{I}_{n_S}^{(1)}\right) \leq \mathrm{var}\left(\hat{I}_{n_S}^{(2)}\right)$ holds for all values of n_S that are larger than $n_S^{\star}$
- Variance reduction is then the search of more efficient estimators

Antithetic variates

- We have:

$$I = \mathbb{E}\left[\varphi\left(X_1, \ldots, X_n\right)\right] = \mathbb{E}\left[Y\right]$$

  where $Y = \varphi\left(X_1, \ldots, X_n\right)$ is a one-dimensional random variable

- It follows that:

$$\hat{I}_{n_S} = \bar{Y}_{n_S} = \frac{1}{n_S} \sum_{s=1}^{n_S} Y_s$$

- We now consider the estimators $\bar{Y}_{n_S}$ and $\bar{Y}'_{n_S}$ based on two different samples and define $\bar{Y}^{\star}$ as follows:

$$\bar{Y}^{\star} = \frac{\bar{Y}_{n_S} + \bar{Y}'_{n_S}}{2}$$

Antithetic variates

- We have:

$$\mathbb{E}\left[\bar{Y}^{\star}\right]=\mathbb{E}\left[\frac{\bar{Y}_{n_S}+\bar{Y}'_{n_S}}{2}\right]=\mathbb{E}\left[\bar{Y}_{n_S}\right]=I$$

and:

$$\begin{aligned}
\operatorname{var}\left(\bar{Y}^{\star}\right) &= \operatorname{var}\left(\frac{\bar{Y}_{n_S}+\bar{Y}'_{n_S}}{2}\right) \\
&= \frac{1}{4}\operatorname{var}\left(\bar{Y}_{n_S}\right)+\frac{1}{4}\operatorname{var}\left(\bar{Y}'_{n_S}\right)+\frac{1}{2}\operatorname{cov}\left(\bar{Y}_{n_S},\bar{Y}'_{n_S}\right) \\
&= \frac{1+\rho\left\langle\bar{Y}_{n_S},\bar{Y}'_{n_S}\right\rangle}{2}\operatorname{var}\left(\bar{Y}_{n_S}\right) \\
&= \frac{1+\rho\left\langle Y_s,Y'_s\right\rangle}{2}\operatorname{var}\left(\bar{Y}_{n_S}\right)
\end{aligned}$$

where $\rho\left\langle Y_s,Y'_s\right\rangle$ is the correlation between Y_s and Y'_s

Antithetic variates

- Because we have $\rho \langle Y_s, Y'_s \rangle \leq 1$, we deduce that:

$$\operatorname{var}\left(\bar{Y}^{\star}\right) \leq \operatorname{var}\left(\bar{Y}_{n_S}\right)$$

- If we simulate the random variates Y_s and Y'_s independently, $\rho \langle Y_s, Y'_s \rangle$ is equal to zero and the variance of the estimator is divided by 2
- However, the number of simulations have been multiplied by two. The efficiency of the estimator has then not been improved

Antithetic variates

- The underlying idea of antithetic variables is therefore to use two perfectly dependent random variables Y_s and Y_s':

$$Y_s' = \psi(Y_s)$$

where ψ is a deterministic function

- This implies that:

$$\bar{Y}^{\star}_{n_S} = \frac{1}{n_S} \sum_{s=1}^{n_S} Y^{\star}_s$$

where:

$$Y^{\star}_s = \frac{Y_s + Y_s'}{2} = \frac{Y_s + \psi(Y_s)}{2}$$

- It follows that:

$$\rho \left\langle \bar{Y}_{n_S}, \bar{Y}'_{n_S} \right\rangle = \rho \left\langle Y, Y' \right\rangle = \rho \left\langle Y, \psi(Y) \right\rangle$$

Antithetic variates

- Minimizing the variance $\operatorname{var}\left(\bar{Y}^{\star}\right)$ is then equivalent to minimize the correlation $\rho\left\langle Y,\psi\left(Y\right)\right\rangle$
- We also know that the correlation reaches its lower bound if the dependence function between Y and $\psi\left(Y\right)$ is equal to the lower Fréchet copula:
$$\mathbf{C}\left\langle Y,\psi\left(Y\right)\right\rangle = \mathbf{C}^{-}$$
- However, $\rho\left\langle Y,\psi\left(Y\right)\right\rangle$ is not necessarily equal to -1 except in some special cases

Antithetic variates

- We consider the one-dimensional case with $Y = \varphi(X)$
- If we assume that φ is an increasing function, it follows that:

$$\mathbf{C}\langle Y, \psi(Y)\rangle = \mathbf{C}\langle \varphi(X), \psi(\varphi(X))\rangle = \mathbf{C}\langle X, \psi(X)\rangle$$

- To obtain the lower bound $\mathbf{C}^{-}$, X and $\psi(X)$ must be countermonotonic:

$$\psi(X) = \mathbf{F}^{-1}(1 - \mathbf{F}(X))$$

  where $\mathbf{F}$ is the probability distribution of X
- For instance, if $X \sim \mathcal{U}_{[0,1]}$, we have $X' = 1 - X$. In the case where $X \sim \mathcal{N}(0,1)$, we have:

$$X' = \Phi^{-1}(1 - \Phi(X)) = \Phi^{-1}(\Phi(-X)) = -X$$

Antithetic variates

Example #9

We consider the following functions:

1. $\varphi_1(x) = x^3 + x + 1$
2. $\varphi_2(x) = x^4 + x^2 + 1$
3. $\varphi_3(x) = x^4 + x^3 + x^2 + x + 1$

Antithetic variates

For each function, we want to estimate $I = \mathbb{E}[\varphi(\mathcal{N}(0,1))]$ using the antithetic estimator:

$$\bar{Y}^{\star}_{n_S} = \frac{1}{n_S}\sum_{s=1}^{n_S}\frac{\varphi(X_s)+\varphi(-X_s)}{2}$$

where $X_s \sim \mathcal{N}(0,1)$

- Let $X \sim \mathcal{N}(0,1)$. We have $\mathbb{E}\left[X^2\right] = 1$, $\mathbb{E}\left[X^{2m}\right] = (2m-1)\,\mathbb{E}\left[X^{2m-2}\right]$ and $\mathbb{E}\left[X^{2m+1}\right] = 0$ for $m \in \mathbb{N}$
- We obtain the following results:

$\varphi(x)$	$\varphi_1(x)$	$\varphi_2(x)$	$\varphi_3(x)$
$\mathbb{E}[\varphi(X_s)]$ or $\mathbb{E}[\varphi(-X_s)]$	1	5	5
$\mathrm{var}(\varphi(X_s))$ or $\mathrm{var}(\varphi(-X_s))$	22	122	144
$\mathrm{cov}(\varphi(X_s),\varphi(-X_s))$	-22	122	100
$\rho\langle\varphi(X_s),\varphi(-X_s)\rangle$	-1	1	$25/36$

Antithetic variates

To understand these numerical results, we must study the relationship between $\mathbf{C}\left\langle X, X'\right\rangle$ and $\mathbf{C}\left\langle Y, Y'\right\rangle$. Indeed, we have:

$$\left\{\mathbf{C}\left\langle X, X'\right\rangle = \mathbf{C}^{-} \Rightarrow \mathbf{C}\left\langle Y, Y'\right\rangle = \mathbf{C}^{-}\right\} \Leftrightarrow \varphi'\left(x\right) \geq 0$$

Antithetic variates

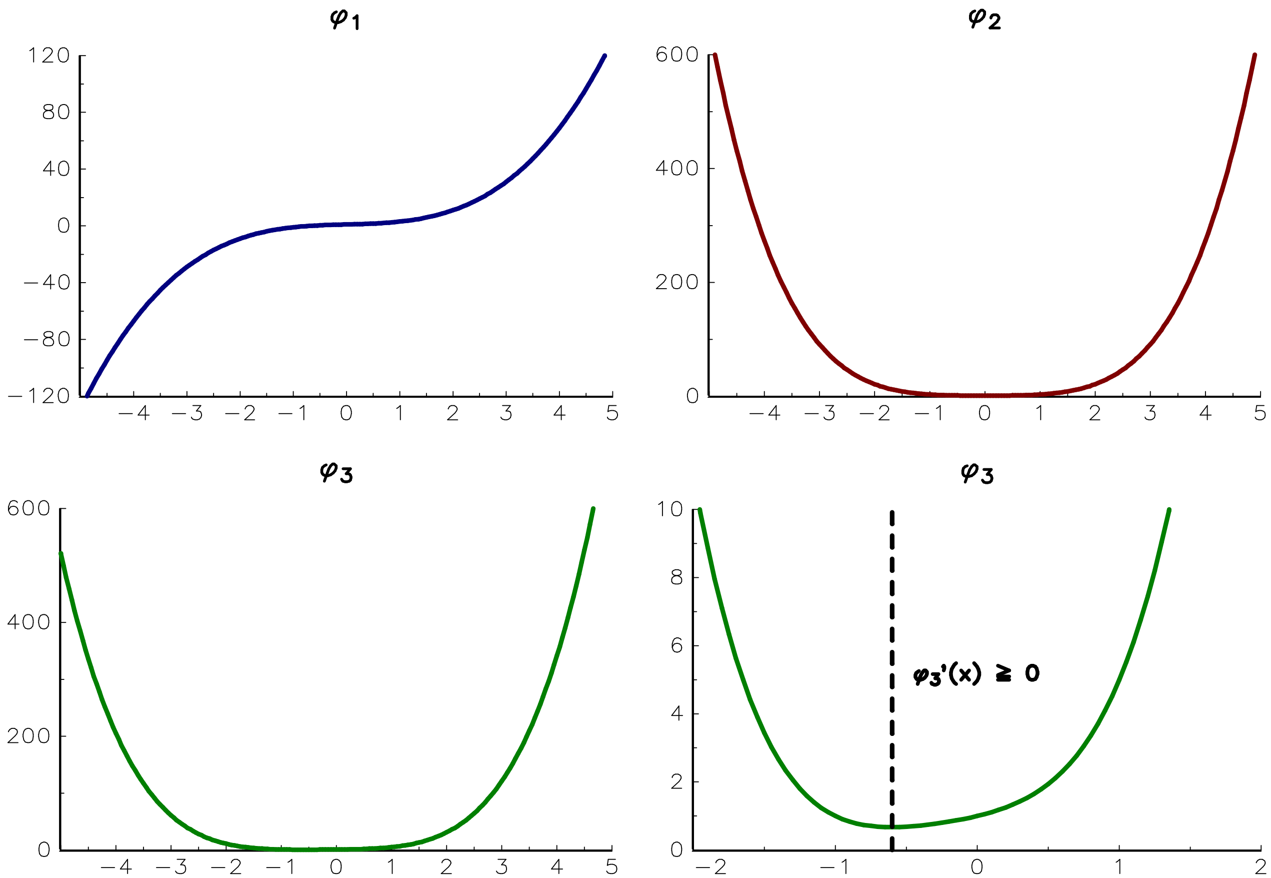


Figure: Functions $\varphi_1(x)$, $\varphi_2(x)$ and $\varphi_3(x)$

Application to the geometric Brownian motion

- In the Gaussian case $X \sim \mathcal{N}(0,1)$, the antithetic variable is:

$$X' = -X$$

- As the simulation of $Y \sim \mathcal{N}\left(\mu, \sigma^2\right)$ is obtained using the relationship $Y = \mu + \sigma X$, we deduce that the antithetic variable is:

$$Y' = \mu - \sigma X = \mu - \sigma \frac{(Y - \mu)}{\sigma} = 2\mu - Y$$

- If we consider the geometric Brownian motion, the fixed-interval scheme is:

$$X_{m+1} = X_m \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) h + \sigma\sqrt{h} \cdot \varepsilon_m\right)$$

whereas the antithetic path is given by:

$$X'_{m+1} = X'_m \cdot \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) h - \sigma\sqrt{h} \cdot \varepsilon_m\right)$$

Application to the geometric Brownian motion

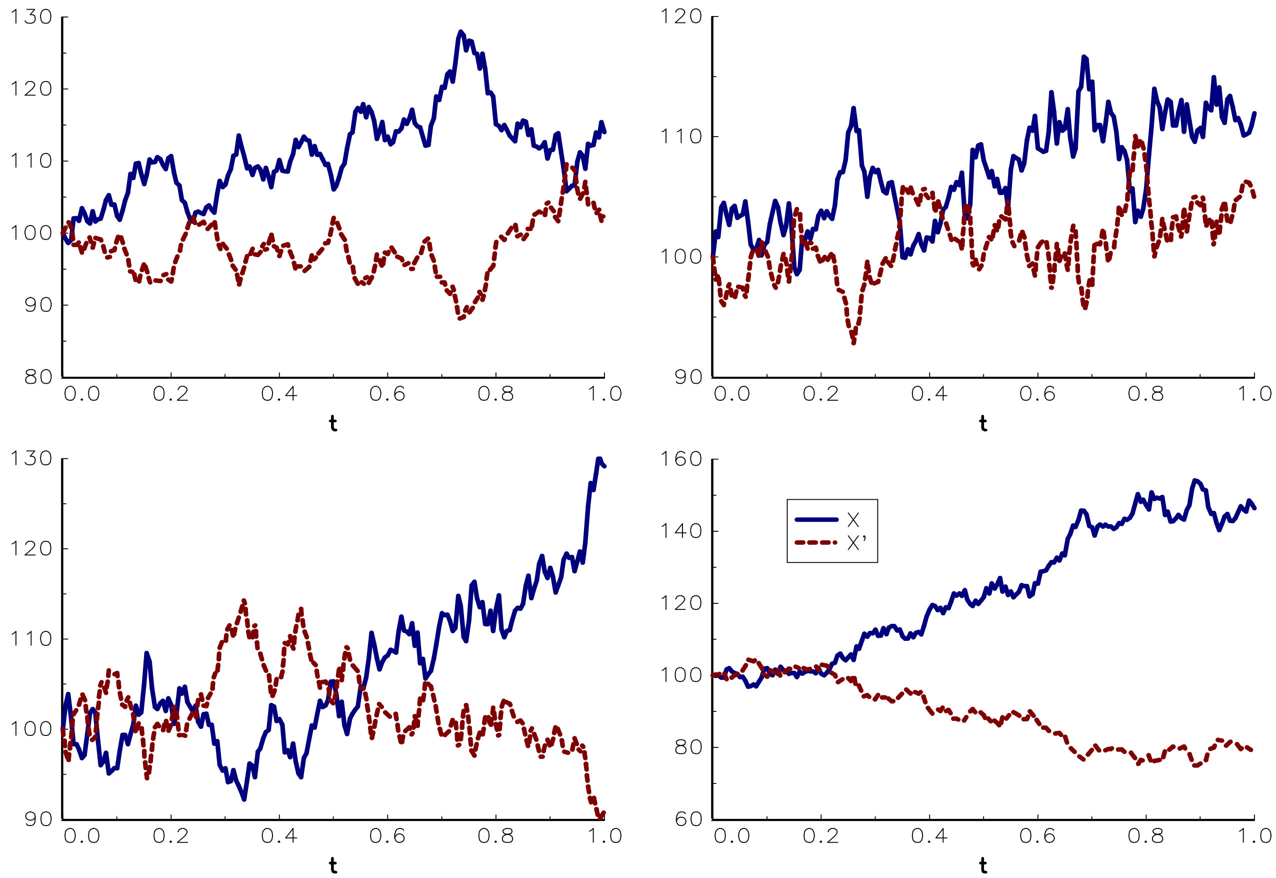


Figure: Antithetic simulation of the GBM process

Application to the geometric Brownian motion

- In the multidimensional case, we recall that:

$$X_{j,m+1} = X_{j,m} \cdot \exp\left(\left(\mu_j - \frac{1}{2}\sigma_j^2\right) h + \sigma_j \sqrt{h} \cdot \varepsilon_{j,m}\right)$$

  where $\varepsilon_m = (\varepsilon_{1,m}, \ldots, \varepsilon_{n,m}) \sim \mathcal{N}_n(\mathbf{0}, \rho)$
- We simulate ε_m by using the relationship $\varepsilon_m = P \cdot \eta_m$ where $\eta_m \sim \mathcal{N}_n(\mathbf{0}, I_n)$ and P is the Cholesky matrix satisfying $PP^\top = \rho$
- The antithetic trajectory is then:

$$X'_{j,m+1} = X'_{j,m} \cdot \exp\left(\left(\mu_j - \frac{1}{2}\sigma_j^2\right) h + \sigma_j \sqrt{k} \cdot \varepsilon'_{j,m}\right)$$

  where:

$$\varepsilon'_m = -P \cdot \eta_m = -\varepsilon_m$$

- We verify that $\varepsilon'_m = (\varepsilon'_{1,m}, \ldots, \varepsilon'_{n,m}) \sim \mathcal{N}_n(\mathbf{0}, \rho)$

Application to the geometric Brownian motion

In the Black-Scholes model, the price of the spread option with maturity T and strike K is given by:

$$\mathcal{C} = e^{-rT}\mathbb{E}\left[\left(S_1(T) - S_2(T) - K\right)^+\right]$$

where the prices $S_1(t)$ and $S_2(t)$ of the underlying assets are given by the following SDE:

$$\left\{\begin{array}{l} \mathrm{d}S_1(t) = rS_1(t)\,\mathrm{d}t + \sigma_1 S_1(t)\,\mathrm{d}W_1(t) \\ \mathrm{d}S_2(t) = rS_2(t)\,\mathrm{d}t + \sigma_2 S_2(t)\,\mathrm{d}W_2(t) \end{array}\right.$$

and $\mathbb{E}\left[W_1(t)\,W_2(t)\right] = \rho\,t$

Application to the geometric Brownian motion

- To calculate the option price using Monte Carlo methods, we simulate the bivariate GBM $S_1(t)$ and $S_2(t)$ and the MC estimator is:

$$\hat{\mathcal{C}}_{\mathrm{MC}} = \frac{e^{-rT}}{n_S} \sum_{s=1}^{n_S} \left(S_1^{(s)}(T) - S_2^{(s)}(T) - K \right)^+$$

where $S_j^{(s)}(T)$ is the s^{th} simulation of the terminal value $S_j(T)$

- For the AV estimator, we obtain:

$$\hat{\mathcal{C}}_{\mathrm{AV}} = \frac{e^{-rT}}{n_S} \sum_{s=1}^{n_S} \frac{\left(S_1^{(s)}(T) - S_2^{(s)}(T) - K \right)^+ + \left(S_1'^{(s)}(T) - S_2'^{(s)}(T) - K \right)^+}{2}$$

where $S_j'^{(s)}(T)$ is the antithetic variate of $S_j^{(s)}(T)$

Application to the geometric Brownian motion

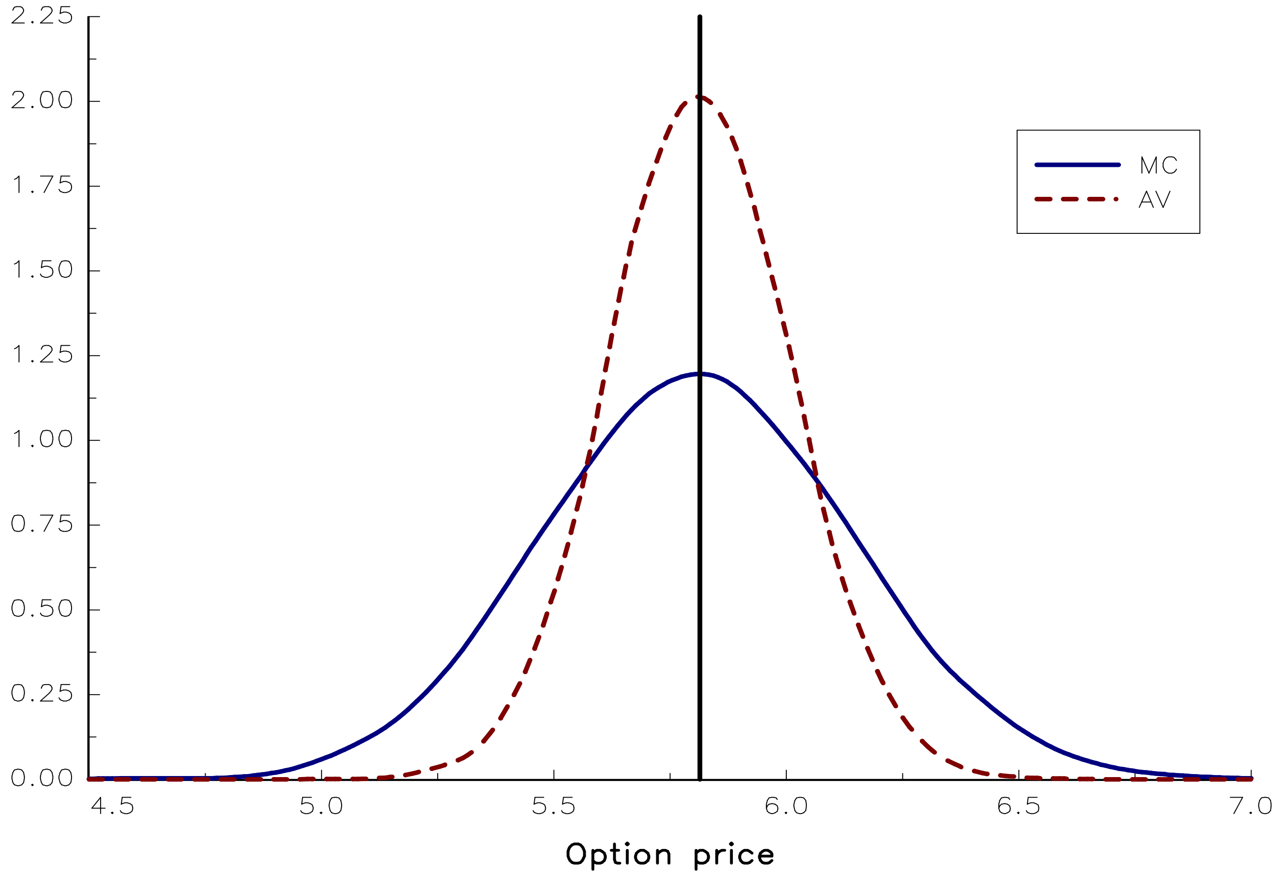


Figure: Probability density function of $\widehat{\mathcal{C}}_{\mathrm{MC}}$ and $\widehat{\mathcal{C}}_{\mathrm{AV}}$ ($n_S = 1\,000$)

Control variates

- Let $Y = \varphi(X_1, \ldots, X_n)$ and V be a random variable with known mean $\mathbb{E}[V]$
- We define Z as follows: $Z = Y + c \cdot (V - \mathbb{E}[V])$
- We deduce that:

$$\begin{aligned}
\mathbb{E}[Z] &= \mathbb{E}[Y + c \cdot (V - \mathbb{E}[V])] \\
&= \mathbb{E}[Y] + c \cdot \mathbb{E}[V - \mathbb{E}[V]] \\
&= \mathbb{E}[\varphi(X_1, \ldots, X_n)]
\end{aligned}$$

and:

$$\begin{aligned}
\operatorname{var}(Z) &= \operatorname{var}(Y + c \cdot (V - \mathbb{E}[V])) \\
&= \operatorname{var}(Y) + 2 \cdot c \cdot \operatorname{cov}(Y, V) + c^2 \cdot \operatorname{var}(V)
\end{aligned}$$

Control variates

- It follows that:

$$\begin{aligned} \operatorname{var}(Z) \leq \operatorname{var}(Y) & \Leftrightarrow 2 \cdot c \cdot \operatorname{cov}(Y, V) + c^2 \cdot \operatorname{var}(V) \leq 0 \\ & \Rightarrow c \cdot \operatorname{cov}(Y, V) \leq 0 \end{aligned}$$

- In order to obtain a lower variance, a necessary condition is that c and $\operatorname{cov}(Y, V)$ have opposite signs
- The minimum is obtained when $\partial_c \operatorname{var}(Z) = 0$ or equivalently when:

$$c^\star = -\frac{\operatorname{cov}(Y, V)}{\operatorname{var}(V)} = -\beta$$

Control variates

- The optimal value $c^{\star}$ is then equal to the opposite of the beta of Y with respect to the control variate V. In this case, we have:

$$Z = Y - \frac{\operatorname{cov}(Y, V)}{\operatorname{var}(V)} \cdot (V - \mathbb{E}[V])$$

and:

$$\operatorname{var}(Z) = \operatorname{var}(Y) - \frac{\operatorname{cov}^2(Y, V)}{\operatorname{var}(V)} = \left(1 - \rho^2 \langle Y, V \rangle\right) \cdot \operatorname{var}(Y)$$

- This implies that we have to choose a control variate V that is highly (positively or negatively) correlated with Y in order to reduce the variance

Control variates

Example

We consider that $X \sim \mathcal{U}_{[0,1]}$ and $\varphi(x) = e^x$. We would like to estimate:

$$I = \mathbb{E}[\varphi(X)] = \int_0^1 e^x \,\mathrm{d}x$$

Control variates

- We set $Y = e^X$ and $V = X$
- We know that $\mathbb{E}[V] = 1/2$ and $\operatorname{var}(V) = 1/12$
- It follows that:

$$
\begin{aligned}
\operatorname{var}(Y) &= \mathbb{E}\left[Y^2\right] - \mathbb{E}^2[Y] \\
&= \int_0^1 e^{2x}\,\mathrm{d}x - \left(\int_0^1 e^x\,\mathrm{d}x\right)^2 \\
&= \left[\frac{e^{2x}}{2}\right]_0^1 - \left(e^1 - e^0\right)^2 \\
&= \frac{4e - e^2 - 3}{2} \\
&\approx 0.2420
\end{aligned}
$$

Control variates

- We have:

$$
\begin{aligned}
\operatorname{cov}(Y, V) &= \mathbb{E}[VY] - \mathbb{E}[V]\,\mathbb{E}[Y] \\
&= \int_0^1 x e^x \,\mathrm{d}x - \frac{1}{2}\left(e^1 - e^0\right) \\
&= \left[x e^x\right]_0^1 - \int_0^1 e^x \,\mathrm{d}x - \frac{1}{2}\left(e^1 - e^0\right) \\
&= \frac{3-e}{2} \\
&\approx 0.1409
\end{aligned}
$$

- If we consider the VC estimator Z defined by:

$$
\begin{aligned}
Z &= Y - \frac{\operatorname{cov}(Y, V)}{\operatorname{var}(V)} \cdot (V - \mathbb{E}[V]) \\
&= Y - (18 - 6e) \cdot \left(V - \frac{1}{2}\right)
\end{aligned}
$$

Control variates

- We have $\beta \approx 1.6903$
- We obtain:

$$\begin{aligned}\operatorname{var}(Z) &= \operatorname{var}(Y) - \frac{\operatorname{cov}^2(Y,V)}{\operatorname{var}(V)} \\ &= \frac{4e - e^2 - 3}{2} - 3 \cdot (3-e)^2 \\ &\approx 0.0039\end{aligned}$$

- We conclude that we have dramatically reduced the variance of the estimator, because we have:

$$\frac{\operatorname{var}\left(\hat{I}_{\mathrm{CV}}\right)}{\operatorname{var}\left(\hat{I}_{\mathrm{MC}}\right)} = \frac{\operatorname{var}(Z)}{\operatorname{var}(Y)} = 1.628\%$$

Control variates

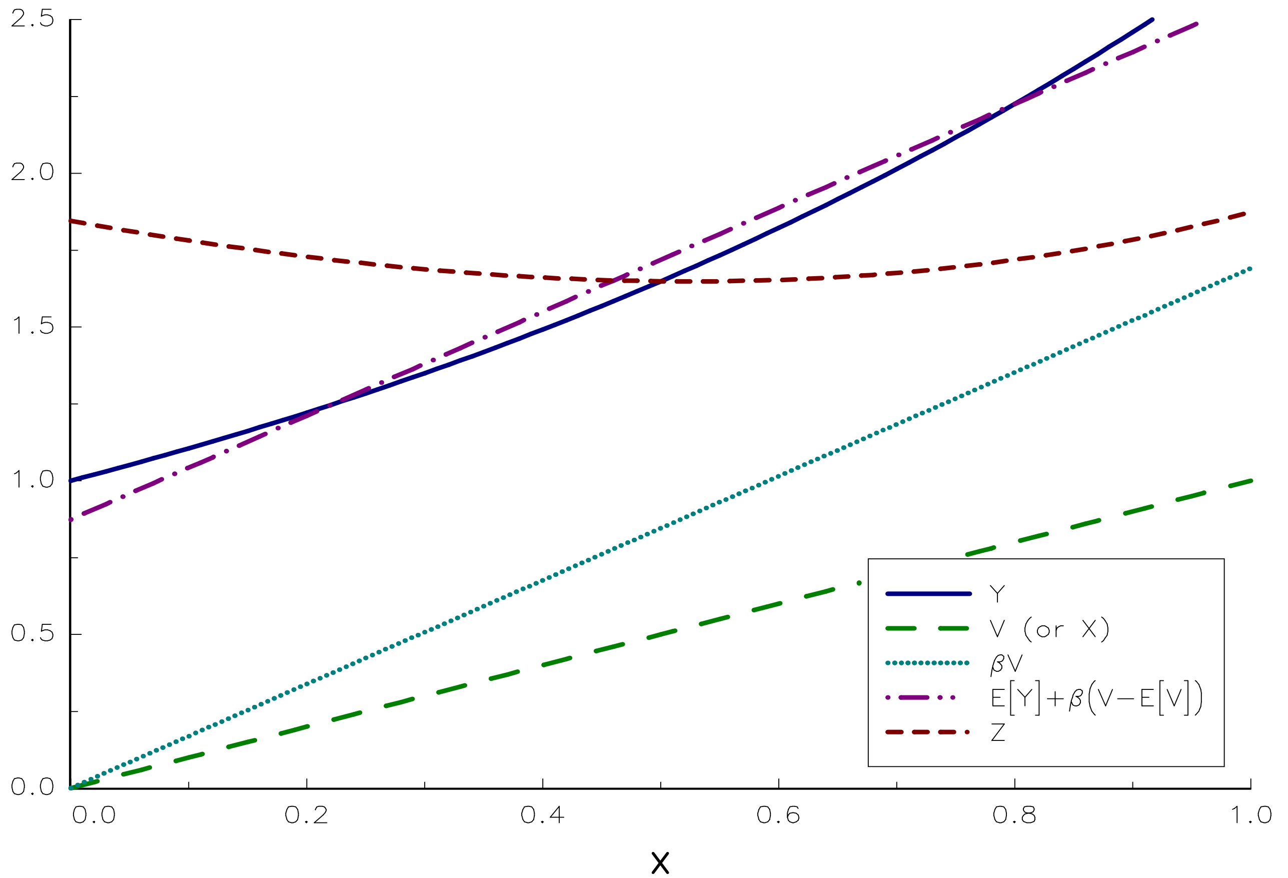


Figure: Understanding the variance reduction in control variates

Control variates

- $\hat{Y}$ is the conditional expectation of Y with respect to V:

$$\mathbb{E}[Y \mid V] = \mathbb{E}[Y] + \beta(V - \mathbf{E}[V])$$

- This is the best linear estimator of Y
- The residual U of the linear regression is then equal to:

$$U = Y - \hat{Y} = (Y - \mathbb{E}[Y]) - \beta(V - \mathbf{E}[V])$$

- The CV estimator Z is a translation of the residual in order to satisfy $\mathbb{E}[Z] = \mathbb{E}[Y]$:

$$Z = \mathbb{E}[Y] + U = Y - \beta(V - \mathbf{E}[V])$$

- By construction, the variance of the residual U is lower than the variance of the random variable Y. We conclude that:

$$\mathrm{var}(Z) = \mathrm{var}(U) \leq \mathrm{var}(Y)$$

Control variates

We can therefore obtain a large variance reduction if the following conditions are satisfied:

- the control variate V largely explains the random variable Y
- the relationship between Y and V is almost linear

Control variates

The price of an arithmetic Asian call option is given by:

$$\mathcal{C} = e^{-rT}\mathbb{E}\left[\left(\bar{S} - K\right)^{+}\right]$$

where K is the strike of the option and $\bar{S}$ denotes the average of $S(t)$ on a given number of fixing dates[21] $\{t_1, \ldots, t_{n_F}\}$:

$$\bar{S} = \frac{1}{n_F}\sum_{m=1}^{n_F} S(t_m)$$

We can estimate the option price using the Black-Scholes model

[21] We have $t_{n_F} = T$.

Control variates

We can also reduce the variance of the MC estimator by considering the following control variates:

1. the terminal value $V_1 = S(T)$ of the underlying asset;
2. the average value $V_2 = \bar{S}$;
3. the discounted payoff of the call option $V_3 = e^{-rT}(S(T) - K)^+$;
4. the discounted payoff of the geometric Asian call option $V_4 = e^{-rT}\left(\tilde{S} - K\right)^+$ where:

$$\tilde{S} = \left(\prod_{m=1}^{n_F} S(t_m)\right)^{1/n_F}$$

Control variates

For these control variates, we know the expected value

- In the first case, we have:

$$\mathbb{E}\left[S\left(T\right)\right] = S_0 e^{rT}$$

- In the first case, we have:

$$\mathbb{E}\left[\bar{S}\right] = \frac{S_0}{n_F} \sum_{m=1}^{n_F} e^{rt_m}$$

Control variates

- The expected value of the third control variate is the Black-Scholes formula of the European call option:

$$\tilde{S} = \left(\prod_{m=1}^{n_F} S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)t_m+\sigma W(t_m)}\right)^{1/n_F} = S_0 \cdot \exp\left(\left(r-\frac{1}{2}\sigma^2\right)\bar{t}+\sigma\bar{W}\right)$$

where:

$$\bar{t} = \frac{1}{n_F}\sum_{m=1}^{n_F} t_m$$

and:

$$\bar{W} = \frac{1}{n_F}\sum_{m=1}^{n_F} W(t_m)$$

- Because $\tilde{S}$ has a log-normal distribution, we deduce that the expected value of the fourth control variate is also given by a Black-Scholes formula

Control variates

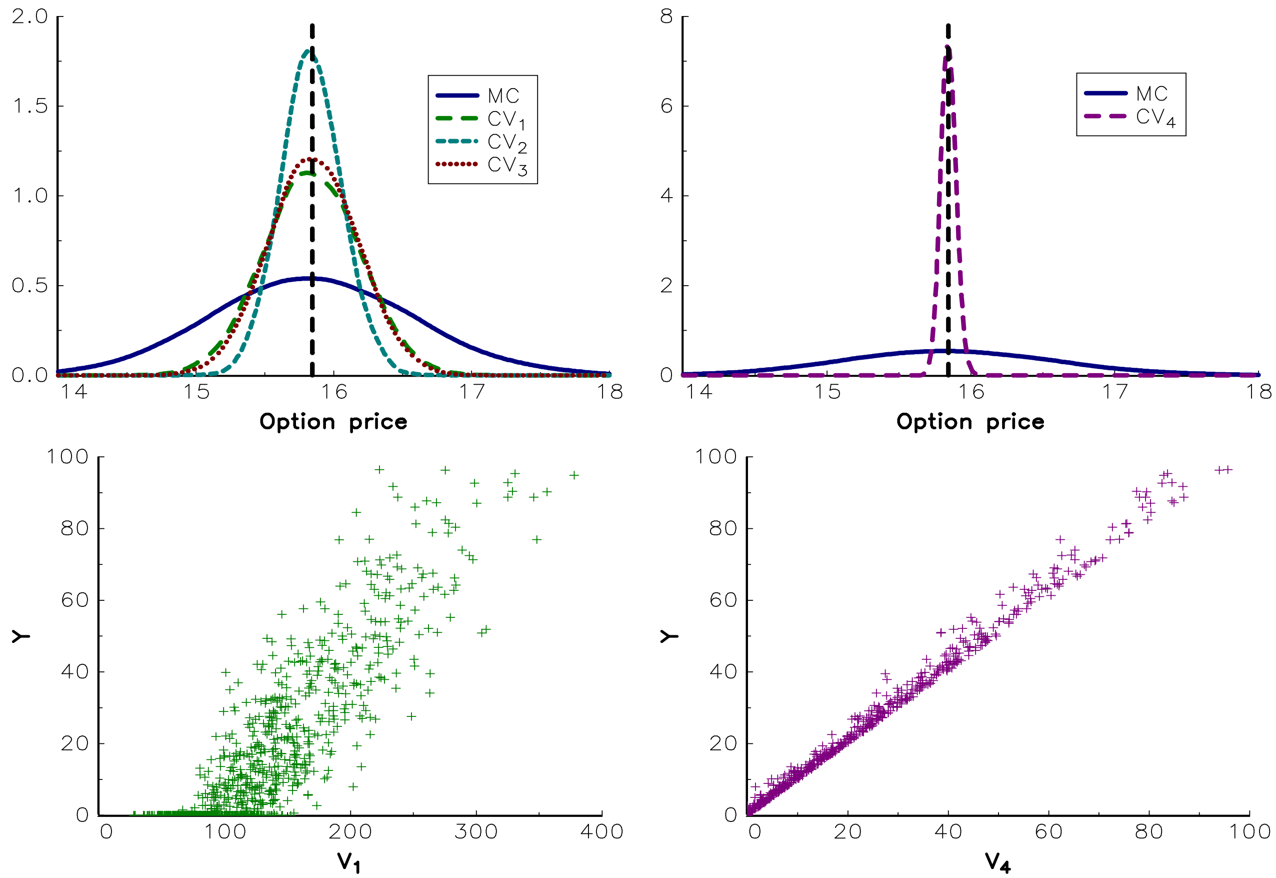


Figure: CV estimator of the arithmetic Asian call option

Control variates

- The previous approach can be extended in the case of several control variates:

$$Z = Y + \sum_{i=1}^{n_{CV}} c_i \cdot (V_i - \mathbb{E}[V_i]) = Y + c^\top (V - \mathbb{E}[V])$$

where $c = (c_1, \ldots, c_{n_{CV}})$ and $V = (V_1, \ldots, V_{n_{CV}})$

- We can show that the optimal value of c is equal to:

$$c^\star = -\operatorname{cov}(V, V)^{-1} \cdot \operatorname{cov}(V, Y)$$

- Minimizing the variance of Z is equivalent to minimize the variance of U:

$$U = Y - \hat{Y} = Y - \left(\alpha + \beta^\top V\right)$$

- We deduce that $c^\star = -\beta$. It follows that

$$\operatorname{var}(Z) = \operatorname{var}(U) = \left(1 - R^2\right) \cdot \operatorname{var}(Y)$$

where R^2 is the R-squared coefficient of the linear regression $Y = \alpha + \beta^\top V + U$

Control variates

Table: Linear regression between the Asian call option and the control variates

$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	R^2	$1 - R^2$
−51.482	0.036	0.538			90.7%	9.3%
−24.025	−0.346	0.595	0.548		96.5%	3.5%
−4.141	0.069		0.410		81.1%	18.9%
−38.727		0.428	0.174		92.9%	7.1%
−1.559	−0.040	0.054	0.111	0.905	99.8%	0.2%

Importance sampling

- Let $X = (X_1, \ldots, X_n)$ be a random vector with distribution function $\mathbf{F}$
- We have:

$$\begin{aligned} I &= \mathbb{E}\left[\varphi\left(X_1, \ldots, X_n\right) \mid \mathbf{F}\right] \\ &= \int \cdots \int \varphi\left(x_1, \ldots, x_n\right) f\left(x_1, \ldots, x_n\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \end{aligned}$$

where $f\left(x_1, \ldots, x_n\right)$ is the probability density function of X

Importance sampling

- It follows that:

$$\begin{aligned} I &= \int \cdots \int \left(\varphi(x_1, \ldots, x_n) \frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)} \right) g(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ &= \mathbb{E}\left[\varphi(X_1, \ldots, X_n) \frac{f(X_1, \ldots, X_n)}{g(X_1, \ldots, X_n)} \middle| \mathbf{G} \right] \\ &= \mathbb{E}\left[\varphi(X_1, \ldots, X_n) \mathcal{L}(X_1, \ldots, X_n) \mid \mathbf{G} \right] \end{aligned}$$

  where $g(x_1, \ldots, x_n)$ is the probability density function of $\mathbf{G}$ and $\mathcal{L}$ is the likelihood ratio:

$$\mathcal{L}(x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)}$$

- The values taken by $\mathcal{L}(x_1, \ldots, x_n)$ are also called the importance sampling weights

Importance sampling

- Using the vector notation, the relationship becomes:

$$\mathbb{E}\left[\varphi(X) \mid \mathbf{F}\right] = \mathbb{E}\left[\varphi(X)\,\mathcal{L}(X) \mid \mathbf{G}\right]$$

- It follows that:

$$\mathbb{E}\left[\hat{I}_{\mathrm{MC}}\right] = \mathbb{E}\left[\hat{I}_{\mathrm{IS}}\right] = I$$

  where $\hat{I}_{\mathrm{MC}}$ and $\hat{I}_{\mathrm{IS}}$ are the Monte Carlo and importance sampling estimators of I

- We also deduce that:

$$\operatorname{var}\left(\hat{I}_{\mathrm{IS}}\right) = \operatorname{var}\left(\varphi(X)\,\mathcal{L}(X) \mid \mathbf{G}\right)$$

Importance sampling

- It follows that:

$$
\begin{aligned}
\operatorname{var}\left(\hat{I}_{\mathrm{IS}}\right) &= \mathbb{E}\left[\varphi^2(X)\,\mathcal{L}^2(X) \mid \mathbf{G}\right] - \mathbb{E}^2\left[\varphi(X)\,\mathcal{L}(X) \mid \mathbf{G}\right] \\
&= \int \varphi^2(x)\,\mathcal{L}^2(x)\,g(x)\,\mathrm{d}x - I^2 \\
&= \int \varphi^2(x)\,\frac{f^2(x)}{g^2(x)}\,g(x)\,\mathrm{d}x - I^2 \\
&= \int \varphi^2(x)\,\frac{f^2(x)}{g(x)}\,\mathrm{d}x - I^2
\end{aligned}
$$

Importance sampling

- If we compare the variance of the two estimators $\hat{I}_{\mathrm{MC}}$ and $\hat{I}_{\mathrm{IS}}$, we obtain:

$$\begin{aligned}
\operatorname{var}\left(\hat{I}_{\mathrm{IS}}\right) - \operatorname{var}\left(\hat{I}_{\mathrm{MC}}\right) &= \int \varphi^2(x)\frac{f^2(x)}{g(x)}\,\mathrm{d}x - \int \varphi^2(x) f(x)\,\mathrm{d}x \\
&= \int \varphi^2(x)\left(\frac{f(x)}{g(x)} - 1\right) f(x)\,\mathrm{d}x \\
&= \int \varphi^2(x)\left(\mathcal{L}(x) - 1\right) f(x)\,\mathrm{d}x
\end{aligned}$$

- The difference may be negative if the weights $\mathcal{L}(x)$ are small ($\mathcal{L}(x) \ll 1$) because the values of $\varphi^2(x) f(x)$ are positive
- The importance sampling approach changes then the importance of some values x by transforming the original probability distribution $\mathbf{F}$ into another probability distribution $\mathbf{G}$

Importance sampling

- The first-order condition is:

$$-\varphi^2(x) \cdot \frac{f^2(x)}{g^2(x)} = \lambda$$

where λ is a constant

- We have:

$$\begin{aligned} g^\star(x) &= \arg\min \operatorname{var}\left(\hat{I}_{\mathrm{IS}}\right) \\ &= \arg\min \int \varphi^2(x) \frac{f^2(x)}{g(x)} \,\mathrm{d}x \\ &= c \cdot |\varphi(x)| \cdot f(x) \end{aligned}$$

where c is the normalizing constant such that $\int g^\star(x) \,\mathrm{d}x = 1$

- A good choice of the IS density $g(x)$ is then an approximation of $|\varphi(x)| \cdot f(x)$ such that $g(x)$ can easily be simulated

Importance sampling

Remark

In order to simplify the notation and avoid confusions, we consider that $X \sim \mathbf{F}$ and $Z \sim \mathbf{G}$ in the sequel. This means that $\hat{I}_{\mathrm{MC}} = \varphi(X)$ and $\hat{I}_{\mathrm{IS}} = \varphi(Z)\,\mathcal{L}(Z)$

Importance sampling

- We consider the estimation of the probability $p = \Pr\{X \geq 3\} \approx 0.1350\%$ when $X \sim \mathcal{N}(0,1)$
- We have:
$$\varphi(x) = \mathbb{1}\{x \geq 3\}$$
- Importance sampling with $Z \sim \mathcal{N}\left(\mu_z, \sigma_z^2\right)$, $\mu_z = 3$ and $\sigma_z = 1 \Rightarrow$ the probability $\Pr\{Z \geq 3\}$ is equal to 50%

Importance sampling

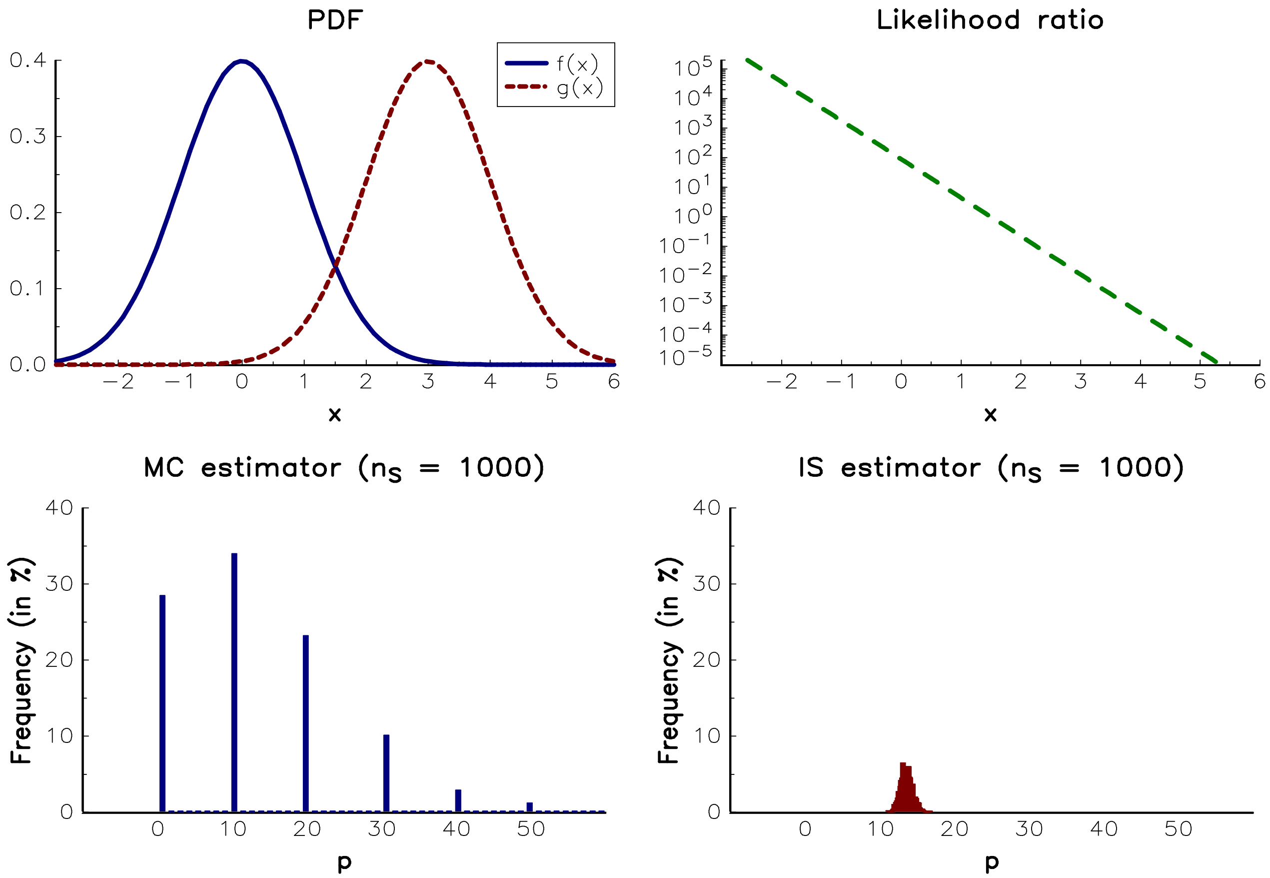


Figure: Histogram of the MC and IS estimators ($n_S = 1\,000$)

Importance sampling

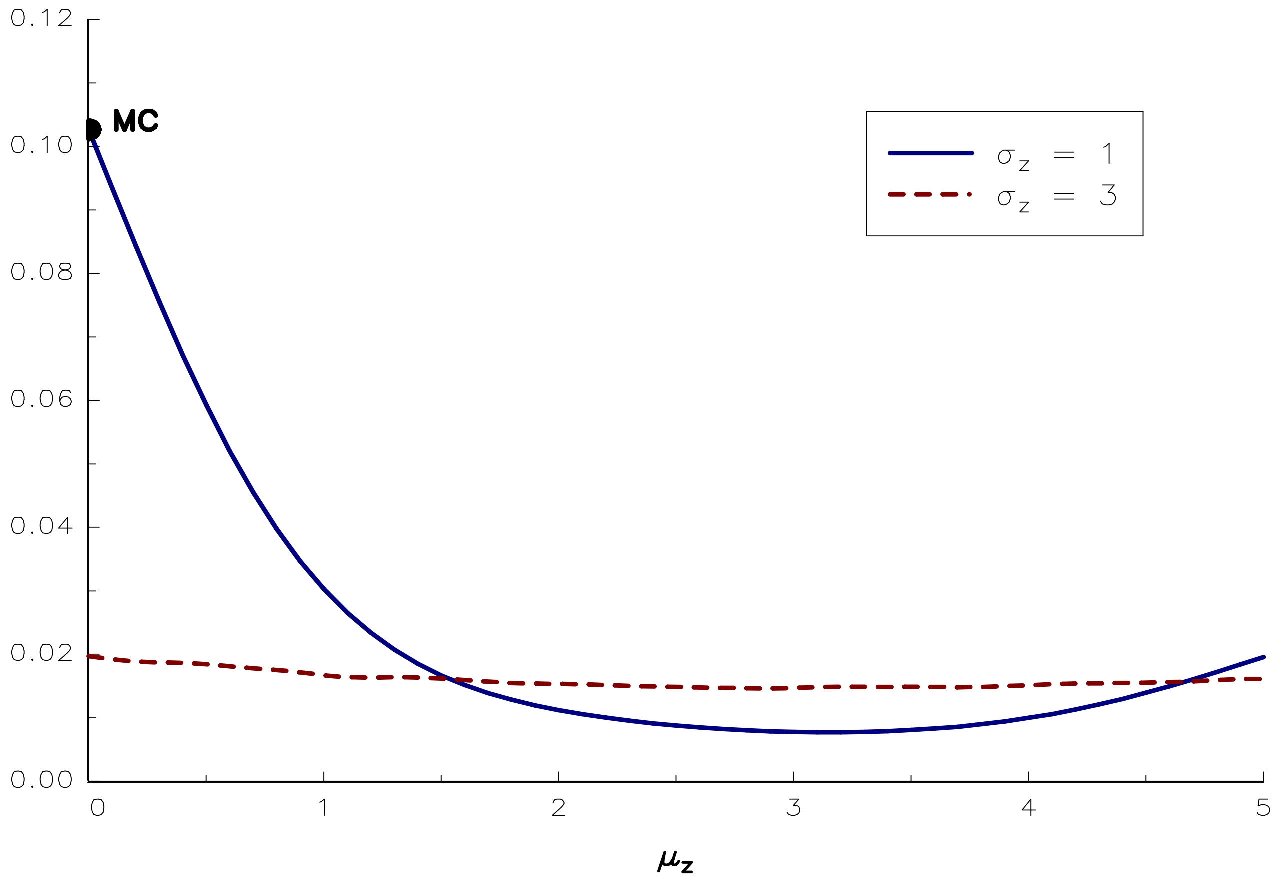


Figure: Standard deviation (in %) of the estimator $\hat{p}_{\mathrm{IS}}$ ($n_S = 1\,000$)

Importance sampling

- We consider the pricing of the put option:

$$\mathcal{P} = e^{-rT}\mathbb{E}\left[(K - S(T))^{+}\right]$$

- We can estimate the option price by using the Monte Carlo method with:

$$\varphi(x) = e^{-rT}(K - x)^{+}$$

- In the case where $K \ll S(0)$, the probability of exercise $\Pr\{S(T) \leq K\}$ is very small
- Therefore, we have to increase the probability of exercise in order to obtain a more efficient estimator

Importance sampling

- In the case of the Black-Scholes model, the density function of $S(T)$ is equal to:

$$f(x) = \frac{1}{x\sigma_x}\phi\left(\frac{\ln x - \mu_x}{\sigma_x}\right)$$

where $\mu_x = \ln S_0 + \left(r - \sigma^2/2\right) T$ and $\sigma_x = \sigma\sqrt{T}$

- We consider the IS density $g(x)$ defined by:

$$g(x) = \frac{1}{x\sigma_z}\phi\left(\frac{\ln x - \mu_z}{\sigma_z}\right)$$

where $\mu_z = \theta + \mu_x$ and $\sigma_z = \sigma_x$

Importance sampling

- For instance, we can choose θ such that the probability of exercise is equal to 50%. It follows that:

$$\begin{aligned}
\Pr\left\{Z \leq K\right\} = \frac{1}{2} & \Leftrightarrow \Phi\left(\frac{\ln K - \theta - \mu_x}{\sigma_x}\right) = \frac{1}{2} \\
& \Leftrightarrow \theta = \ln K - \mu_x \\
& \Leftrightarrow \theta = \ln \frac{K}{S_0} - \left(r - \frac{1}{2}\sigma^2\right) T
\end{aligned}$$

Importance sampling

- We deduce that:

$$\mathcal{P} = \mathbb{E}\left[\varphi\left(S\left(T\right)\right)\right] = \mathbb{E}\left[\varphi\left(S'\left(T\right)\right) \cdot \mathcal{L}\left(S'\left(T\right)\right)\right]$$

where:

$$\mathcal{L}\left(x\right) = \frac{\dfrac{1}{x\sigma_x}\phi\left(\dfrac{\ln x - \mu_x}{\sigma_x}\right)}{\dfrac{1}{x\sigma_z}\phi\left(\dfrac{\ln x - \mu_z}{\sigma_z}\right)} = \exp\left(\frac{\theta^2}{2\sigma_x^2} - \left(\frac{\ln x - \mu_x}{\sigma_x}\right)\cdot\frac{\theta}{\sigma_x}\right)$$

and $S'\left(T\right)$ is the same geometric Brownian motion than $S\left(T\right)$, but with another initial value:

$$S'\left(0\right) = S\left(0\right)e^{\theta} = Ke^{-\left(r-\sigma^2/2\right)T}$$

Importance sampling

Example #10

We assume that $S_0 = 100$, $K = 60$, $r = 5\%$, $\sigma = 20\%$ and $T = 2$. If we consider the previous method, the IS process is simulated using the initial value $S'(0) = Ke^{-\left(r-\sigma^2/2\right)T} = 56.506$, whereas the value of θ is equal to -0.5708

Importance sampling

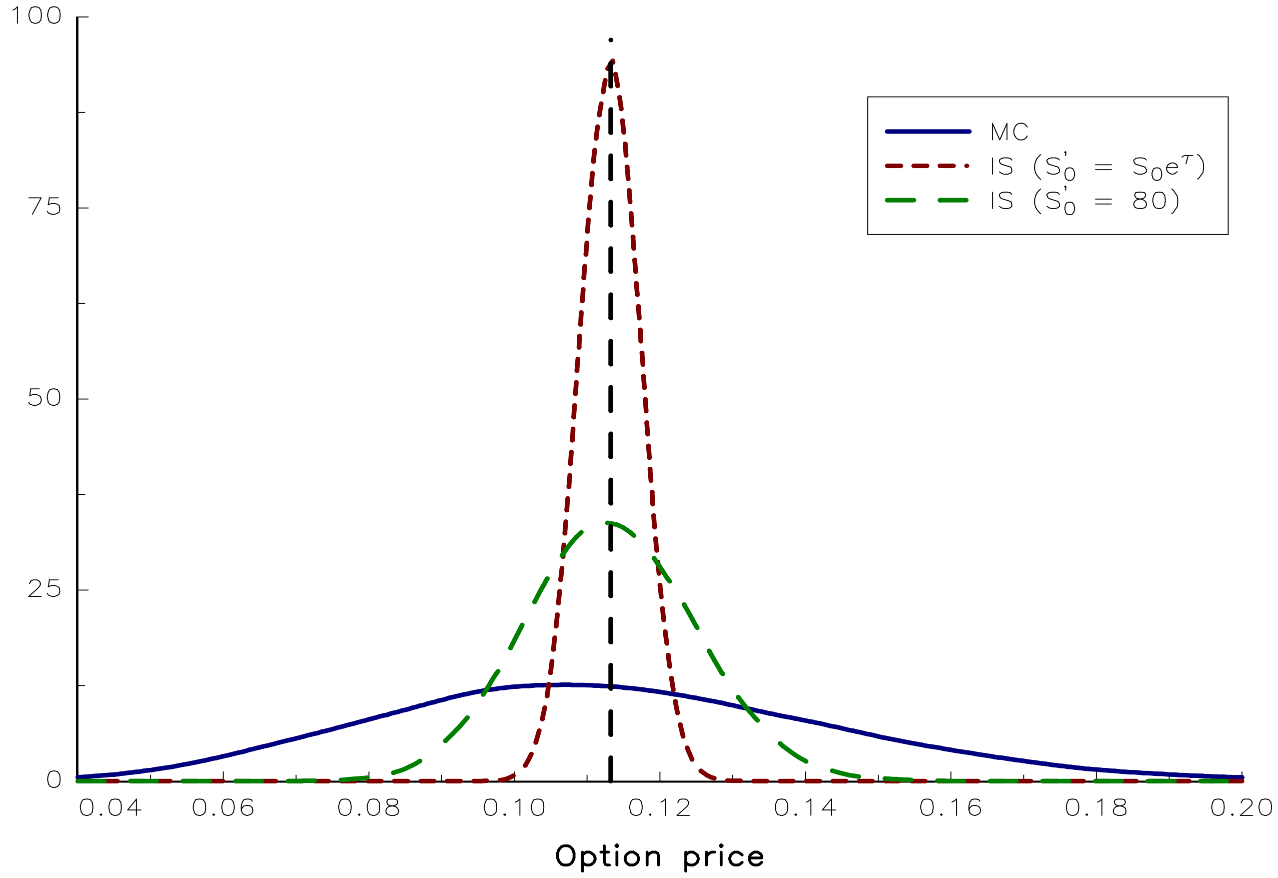


Figure: Density function of the estimators $\hat{\mathcal{P}}_{\mathrm{MC}}$ and $\hat{\mathcal{P}}_{\mathrm{IS}}$ ($n_S = 1\,000$)

Quasi-Monte Carlo simulation methods

- We consider the following Monte Carlo problem:

$$I = \int \cdots \int_{[0,1]^n} \varphi(x_1, \ldots, x_n) \,\mathrm{d}x_1 \cdots \mathrm{d}x_n$$

- Let X be the random vector of independent uniform random variables. It follows that $I = \mathbb{E}[\varphi(X)]$
- The Monte Carlo method consists in generating uniform coordinates in the hypercube $[0, 1]^n$
- Quasi-Monte Carlo methods use non-random coordinates in order to obtain a more nicely uniform distribution

Quasi-Monte Carlo simulation methods

A low discrepancy sequence $\mathcal{U} = \{u_1, \ldots, u_{n_S}\}$ is a set of deterministic points distributed in the hypercube $[0,1]^n$

Quasi-Monte Carlo simulation methods

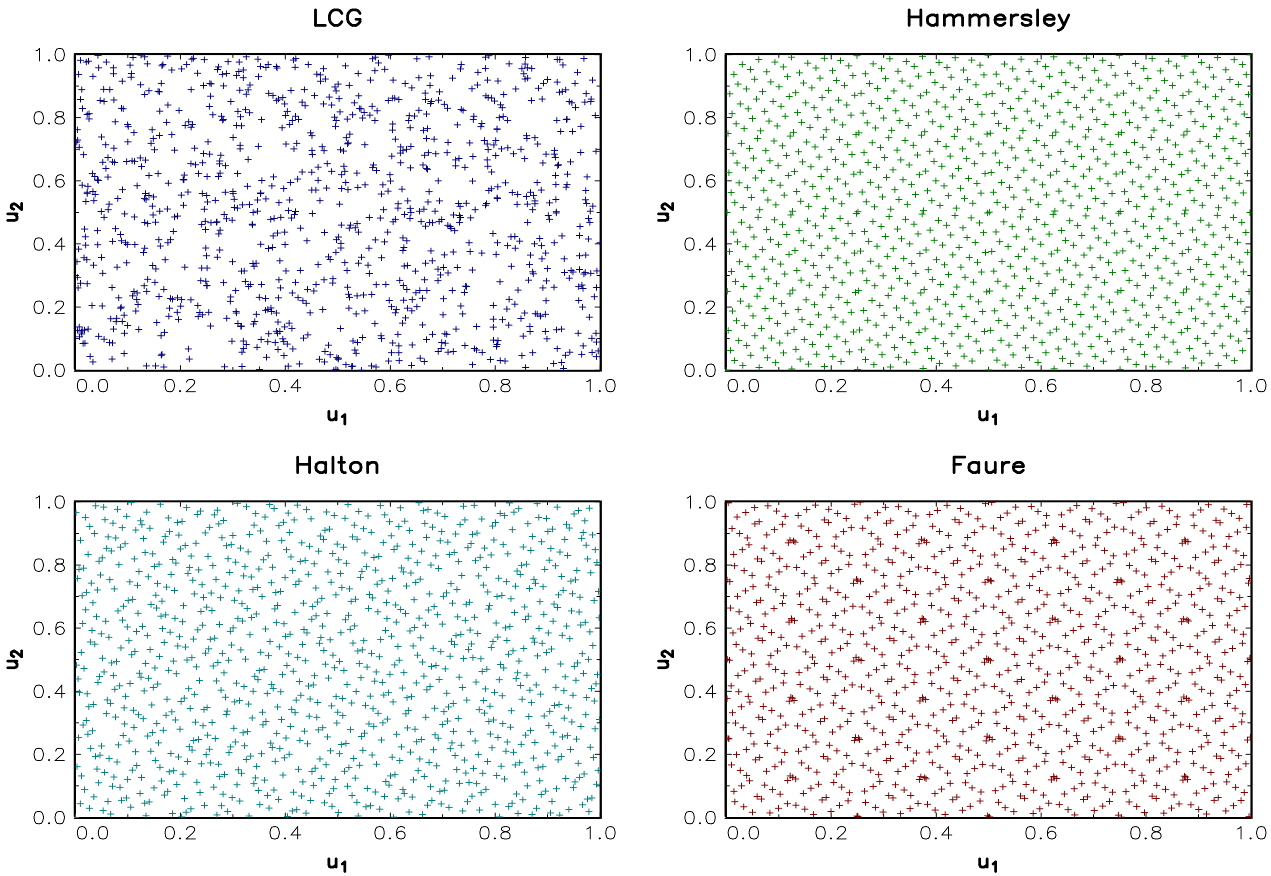


Figure: Comparison of different low discrepancy sequences

Quasi-Monte Carlo simulation methods

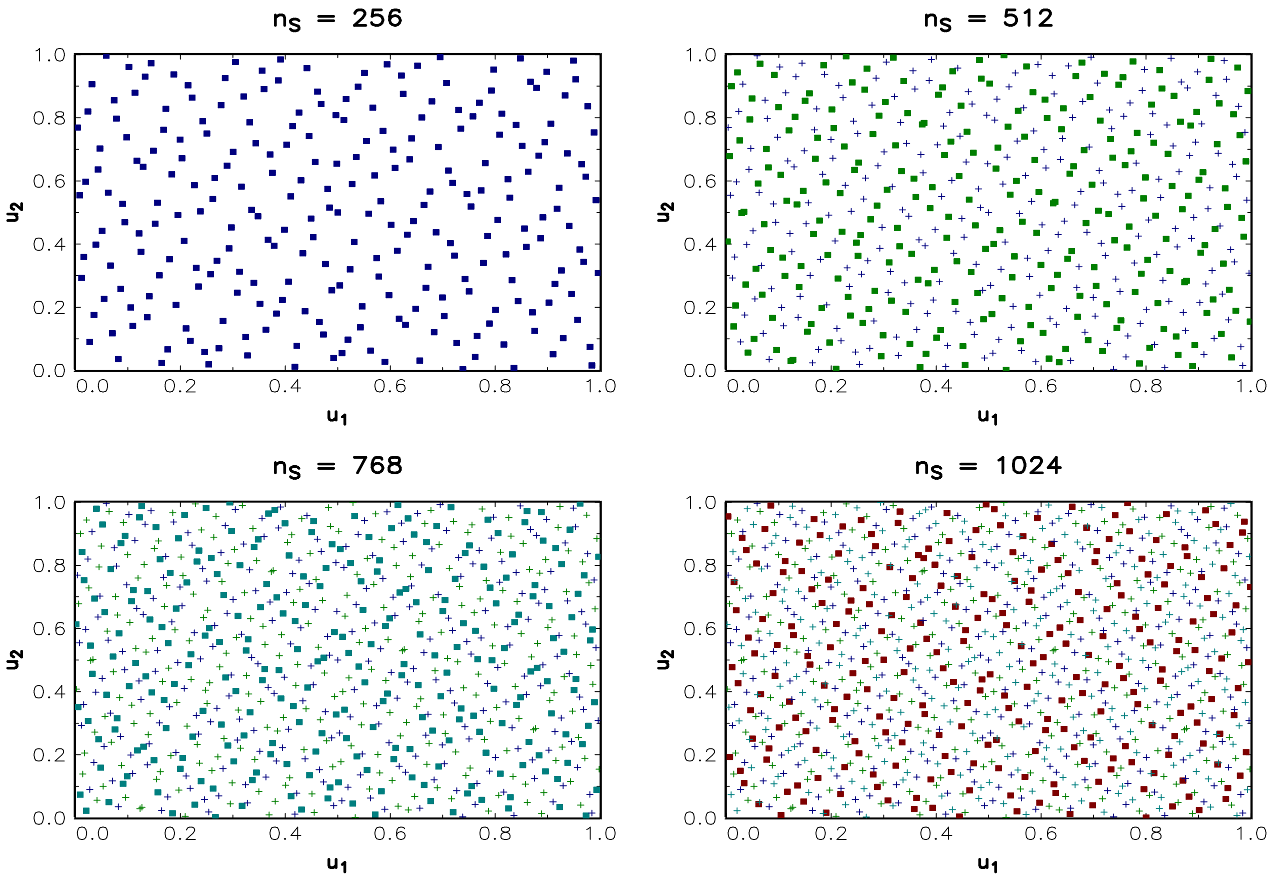


Figure: The Sobol generator

Quasi-Monte Carlo simulation methods

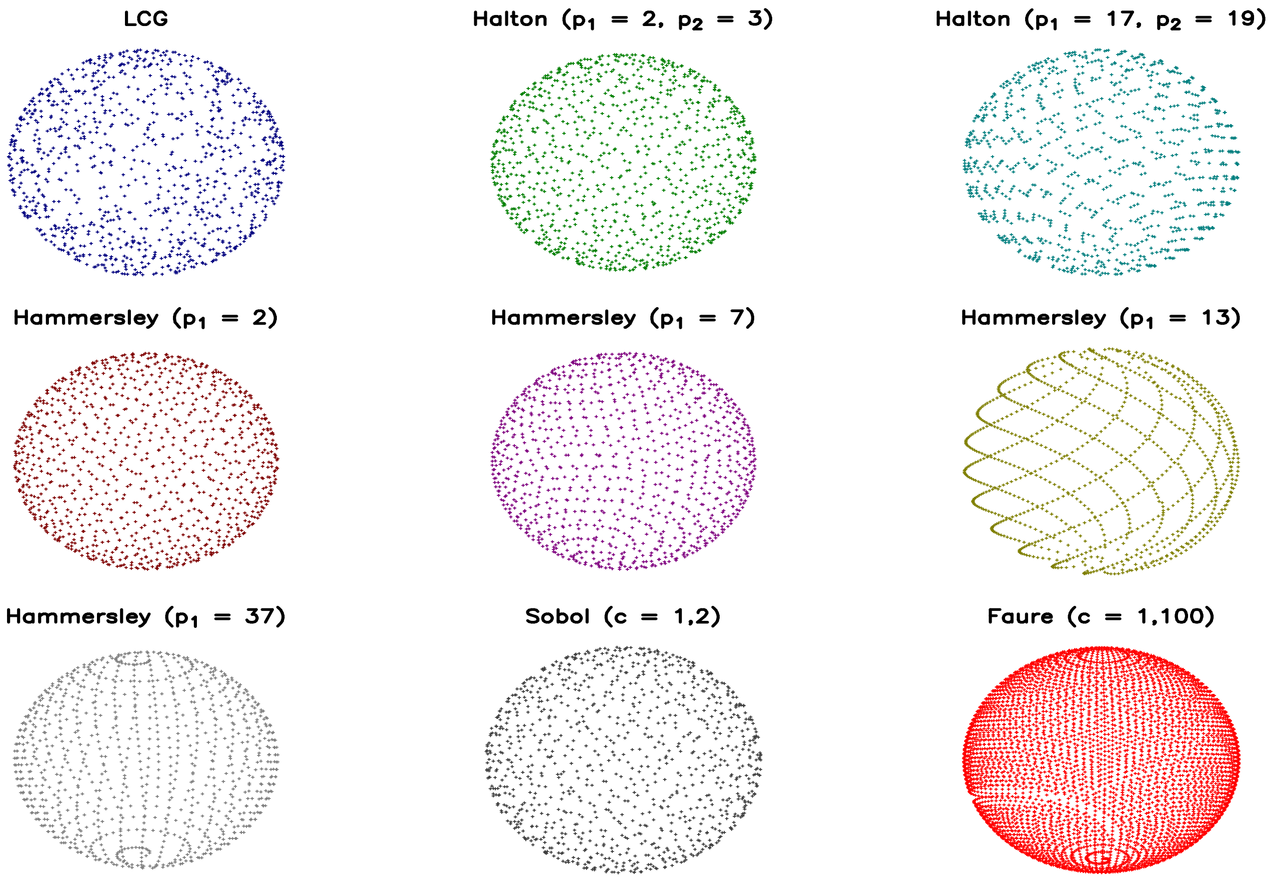


Figure: Quasi-random points on the unit sphere

Quasi-Monte Carlo simulation methods

Example #11

We consider a spread option whose payoff is equal to $(S_1(T) - S_2(T) - K)^+$. The price is calculated using the Black-Scholes model, and the following parameters: $S_1(0) = S_2(0) = 100$, $\sigma_1 = \sigma_2 = 20\%$, $\rho = 50\%$ and $r = 5\%$. The maturity T of the option is set to one year, whereas the strike K is equal to 5. The true price of the spread option is equal to 5.8198.

Quasi-Monte Carlo simulation methods

Table: Pricing of the spread option using quasi-Monte Carlo methods

n_S	10^2	10^3	10^4	10^5	10^6	5×10^6
LCG (1)	4.3988	5.9173	5.8050	5.8326	5.8215	5.8139
LCG (2)	6.1504	6.1640	5.8370	5.8219	5.8265	5.8198
LCG (3)	6.1469	5.7811	5.8125	5.8015	5.8142	5.8197
Hammersley (1)	32.7510	26.5326	21.5500	16.1155	9.0914	5.8199
Hammersley (2)	32.9082	26.4629	21.5465	16.1149	9.0914	5.8199
Halton (1)	8.6256	6.1205	5.8493	5.8228	5.8209	5.8208
Halton (2)	10.6415	6.0526	5.8544	5.8246	5.8208	5.8207
Halton (3)	8.5292	6.0575	5.8474	5.8235	5.8212	5.8208
Sobol	5.7181	5.7598	5.8163	5.8190	5.8198	5.8198
Faure	5.7256	5.7718	5.8157	5.8192	5.8197	5.8198

Exercises

- Exercise 13.4.1 – Simulating random numbers using the inversion method
- Exercise 13.4.6 – Simulation of the bivariate Normal copula
- Exercise 13.4.7 – Computing the capital charge for operational risk

References

DEVROYE, L. (1986)
Non-Uniform Random Variate Generation, Springer-Verlag.

RONCALLI, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 13.

Course 2023-2024 in Financial Risk Management
Lecture 11. Stress Testing and Scenario Analysis

Thierry Roncalli*

*Amundi Asset Management[22]

*University of Paris-Saclay

September 2023

[22]The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- **Lecture 11: Stress Testing and Scenario Analysis**
- Lecture 12: Credit Scoring Models

"Stress testing is now a critical element of risk management for banks and a core tool for banking supervisors and macroprudential authorities" (BCBS, 2017, page 5).

General objective

If we consider a trading book portfolio, we recall that:

$$L_s(w) = P_t(w) - g(\mathcal{F}_{1,s}, \ldots, \mathcal{F}_{m,s}; w)$$

In the case of a stress testing program, we have:

$$L_{\mathrm{stress}}(w) = P_t(w) - g(\mathcal{F}_{1,\mathrm{stress}}, \ldots, \mathcal{F}_{m,\mathrm{stress}}; w)$$

where $(\mathcal{F}_{1,\mathrm{stress}}, \ldots, \mathcal{F}_{m,\mathrm{stress}})$ is the stress scenario

Scenario design and risk factors

2004 FSAP stress scenarios applied to the French banking system

- F_1 flattening of the yield curve due to an increase in interest rates: increase of 150 basis points (bp) in overnight rates, increase of 50 bp in 10-year rates, with interpolation for intermediate maturities
- F_5 share price decline of 30% in all stock markets
- F_9 flattening of the yield curve (increase of 150 basis points in overnight rates, increase of 50 bp in 10-year rates) together with a 30% drop in stock markets
- M_2 increase to USD 40 in the price per barrel of Brent crude for two years (an increase of 48% compared with USD 27 per barrel in the baseline case), without any reaction from the central bank; the increase in the price of oil leads to an increase in the general rate of inflation and a decline in economic activity in France together with a drop in global demand

Scenario design and risk factors

Classification

1. historical scenario: "*a stress test scenario that aims at replicating the changes in risk factor shocks that took place in an actual past episode*"
2. hypothetical scenario: "*a stress test scenario consisting of a hypothetical set of risk factor changes, which does not aim to replicate a historical episode of distress*"
3. macroeconomic scenario: "*a stress test that implements a link between stressed macroeconomic factors [...] and the financial sustainability of either a single financial institution or the entire financial system*"
4. liquidity scenario: "*a liquidity stress test is the process of assessing the impact of an adverse scenario on institution's cash flows as well as on the availability of funding sources, and on market prices of liquid assets*"

Scenario design and risk factors

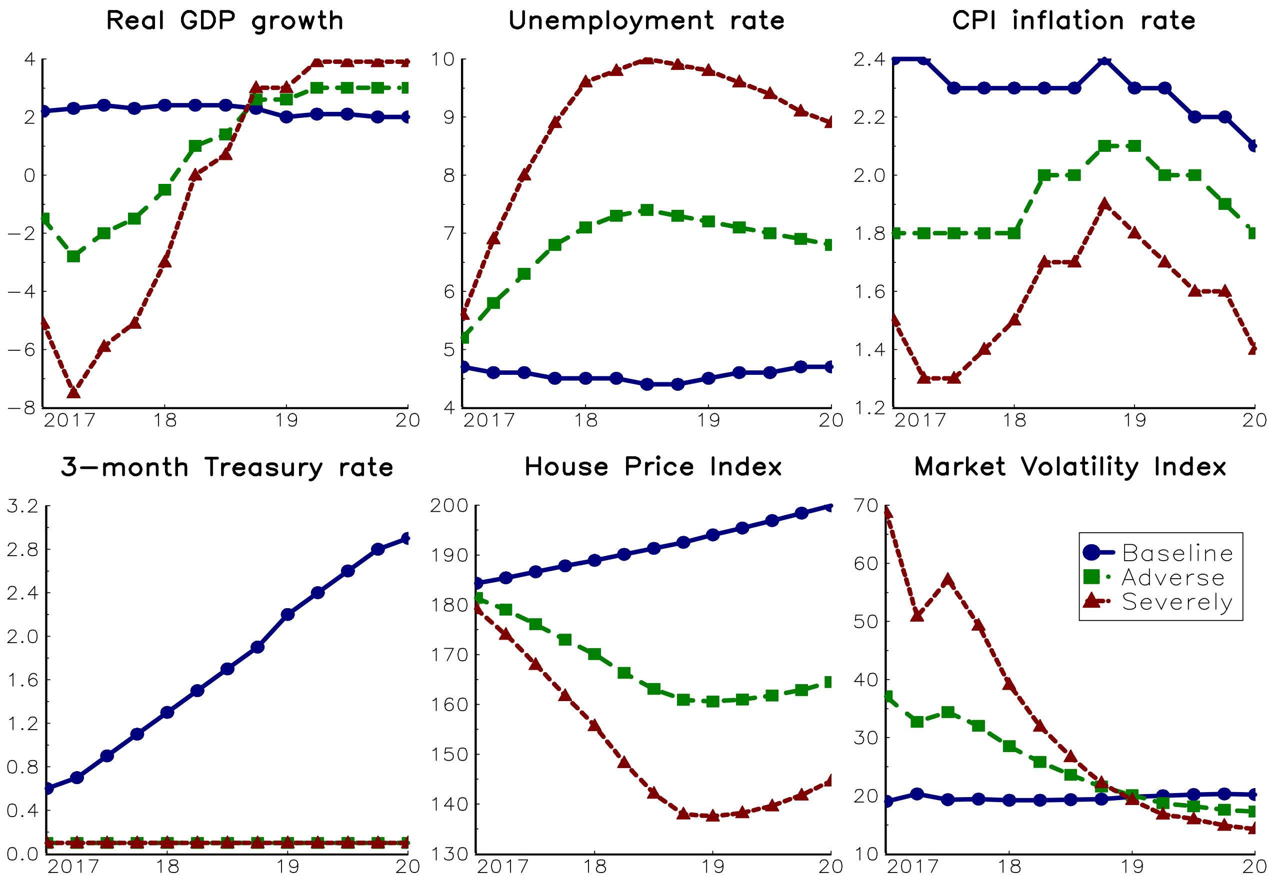


Figure: 2017 DFAST supervisory scenarios: Domestic variables

Scenario design and risk factors

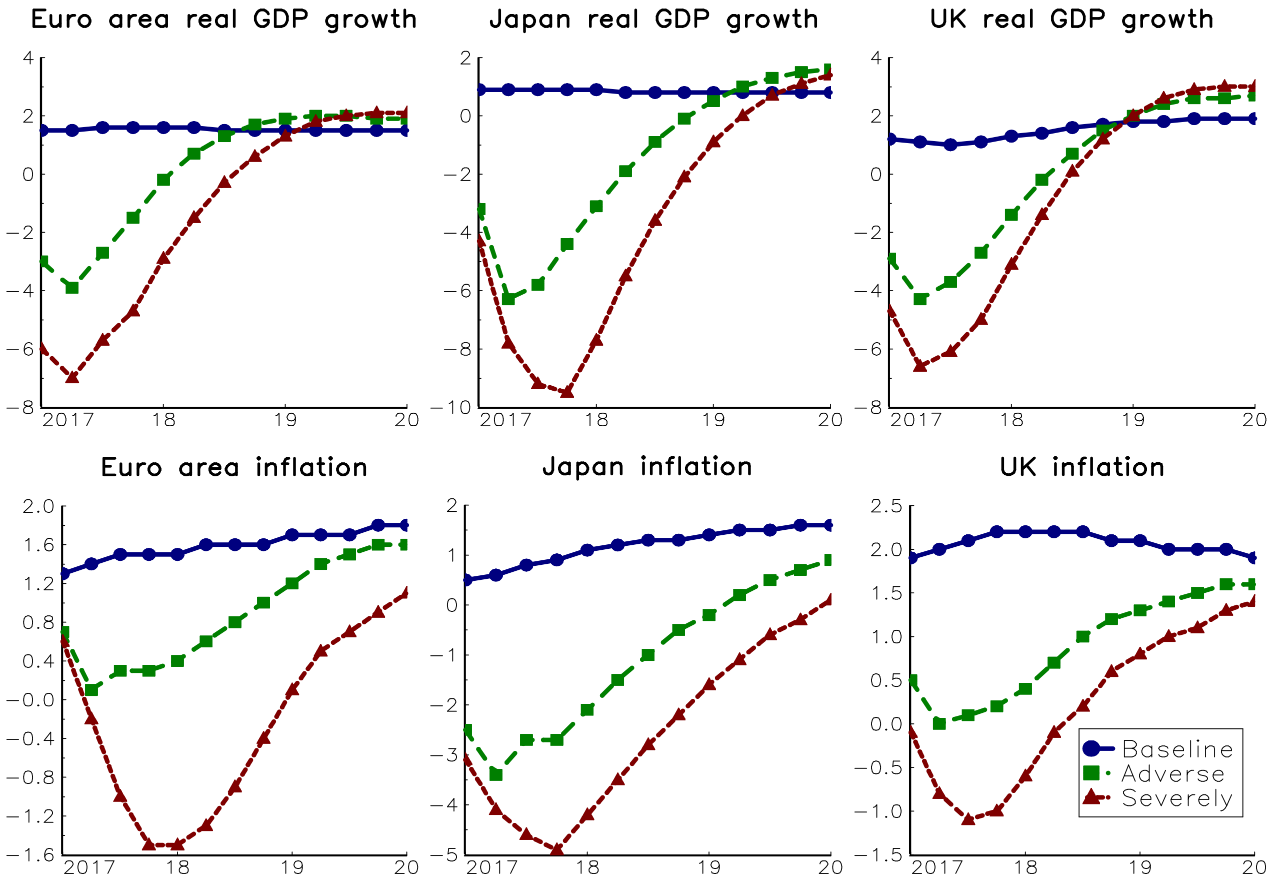


Figure: 2017 DFAST supervisory scenarios: International variables

Firm-specific versus supervisory stress testing

Examples of hard trading limits:

- Unobservable parameters (e.g. correlations of basket options)
- Less liquid assets

Examples of supervisory stress testing:

- Financial sector assessment program (FSAP)
- Dodd-Frank Act stress test (DFAST)
- EU-wide stress testing

Historical approach

Table: Worst historical scenarios of the S&P 500 index

Sc.	1D		1W		1M	
1	1987-10-19	−20.47	1987-10-19	−27.33	2008-10-27	−30.02
2	2008-10-15	−9.03	2008-10-09	−18.34	1987-10-26	−28.89
3	2008-12-01	−8.93	2008-11-20	−17.43	2009-03-09	−22.11
4	2008-09-29	−8.79	2008-10-27	−13.85	2002-07-23	−19.65
5	1987-10-26	−8.28	2011-08-08	−13.01	2001-09-21	−16.89
Sc.	2M		3M		6M	
1	2008-11-20	−37.66	2008-11-20	−41.11	2009-03-09	−46.64
2	1987-10-26	−31.95	1987-11-30	−30.17	1974-09-13	−34.33
3	2002-07-23	−27.29	1974-09-13	−28.59	2002-10-09	−31.29
4	2009-03-06	−26.89	2002-07-23	−27.55	1962-06-27	−26.59
5	1962-06-22	−23.05	2009-03-09	−25.63	1970-05-26	−25.45

Macro-economic approach

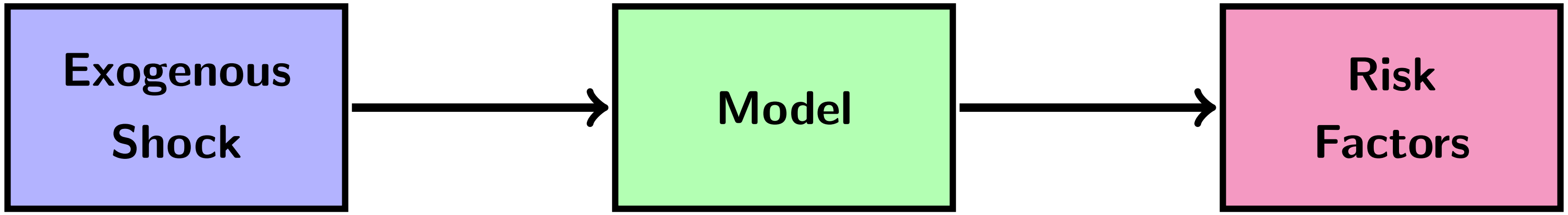


Figure: Macroeconomic approach of stress testing

Macro-economic approach

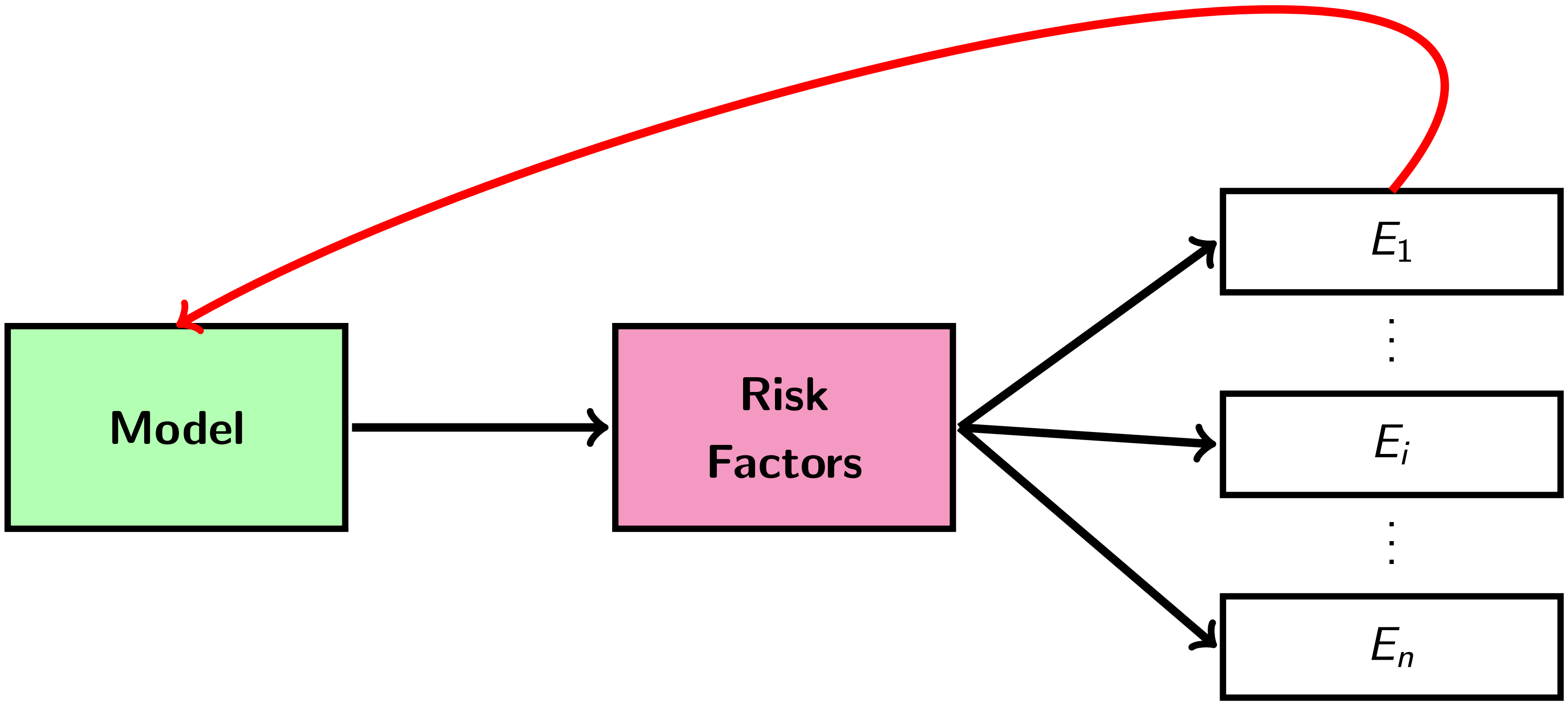


Figure: Feedback effects in stress testing models

Probabilistic approach

At first approximation, a stress scenario can be seen as an extreme quantile or value-at-risk $\Rightarrow$ we can use EVT (extreme value theory)

Univariate stress scenarios

- Let X be the random variable that produces the stress scenario $\mathbb{S}(X)$. If $X \sim \mathbf{F}$ and the relationship between $L(w)$ and X is decreasing, we have:
$$\Pr\{X \leq \mathbb{S}(X)\} = \mathbf{F}(\mathbb{S}(X))$$
- Given a stress scenario $\mathbb{S}(X)$, we deduce its severity:
$$\alpha = \mathbf{F}(\mathbb{S}(X))$$
- We can also compute the stressed value given the probability of occurrence α:
$$\mathbb{S}(X) = \mathbf{F}^{-1}(\alpha)$$

$\alpha \approx 0$ **($\neq$ value-at-risk)**

Univariate stress scenarios

Return time

- We have $\mathcal{T} = \alpha^{-1}$ and $\alpha = \mathcal{T}^{-1}$
- We reiterate that:

$$\mathcal{T} = \alpha^{-1} = n \cdot (1 - \alpha_{\mathrm{GEV}})^{-1}$$

where n is the length of the block maxima

Table: Probability (in %) associated to the return period $\mathcal{T}$ in years

Return period	1	5	10	20	30	50
Daily	0.3846	0.0769	0.0385	0.0192	0.0128	0.0077
Weekly	1.9231	0.3846	0.1923	0.0962	0.0641	0.0385
Monthly	8.3333	1.6667	0.8333	0.4167	0.2778	0.1667
$1 - \alpha_{\mathrm{GEV}}$	7.6923	1.5385	0.7692	0.3846	0.2564	0.1538

Univariate stress scenarios

Table: GEV parameter estimates (in %) of MSCI USA and MSCI EMU indices

Parameter	Long position		Short position	
	MSCI USA	MSCI EMU	MSCI USA	MSCI EMU
μ	1.242	1.572	1.317	1.599
σ	0.720	0.844	0.577	0.730
ξ	19.363	21.603	26.341	26.494

Univariate stress scenarios

Table: Stress scenarios (in %) of MSCI USA and MSCI EMU indices

Year	Long position		Short position	
	MSCI USA	MSCI EMU	MSCI USA	MSCI EMU
5	−5.86	−7.27	5.69	7.16
10	−7.06	−8.83	7.01	8.84
25	−8.92	−11.29	9.17	11.60
50	−10.56	−13.49	11.18	14.17
75	−11.62	−14.94	12.54	15.91
100	−12.43	−16.05	13.59	17.26
Extreme statistic	−9.51	−10.94	11.04	10.87
$\mathcal{T}^{\star}$	32.49	22.24	47.87	20.03

Univariate stress scenarios

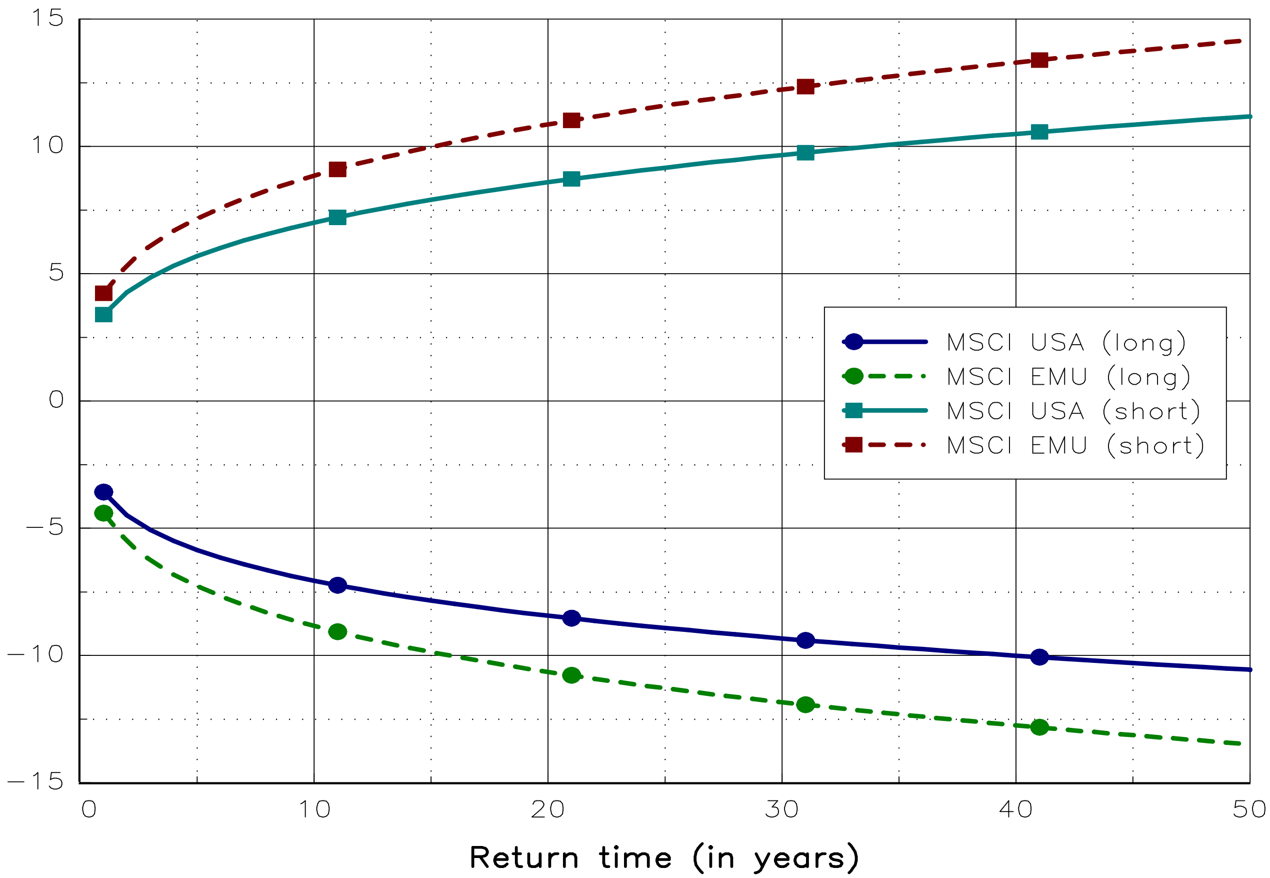


Figure: Stress scenarios (in %) of MSCI USA and MSCI EMU indices

Bivariate stress scenarios

- We note $p = \Pr\left\{X_{n:n,1} > \mathbb{S}\left(X_1\right), X_{n:n,2} > \mathbb{S}\left(X_2\right)\right\}$ the joint probability of stress scenarios $\left(\mathbb{S}\left(X_1\right), \mathbb{S}\left(X_2\right)\right)$
- We have:

$$\begin{aligned} p &= 1 - \mathbf{F}_1\left(\mathbb{S}\left(X_1\right)\right) - \mathbf{F}_2\left(\mathbb{S}\left(X_2\right)\right) + \mathbf{C}\left(\mathbf{F}_1\left(\mathbb{S}\left(X_1\right)\right), \mathbf{F}_2\left(\mathbb{S}\left(X_2\right)\right)\right) \\ &= \bar{\mathbf{C}}\left(\mathbf{F}_1\left(\mathbb{S}\left(X_1\right)\right), \mathbf{F}_2\left(\mathbb{S}\left(X_2\right)\right)\right) \end{aligned}$$

  where $\bar{\mathbf{C}}\left(u_1, u_2\right) = 1 - u_1 - u_2 + \mathbf{C}\left(u_1, u_2\right)$
- We deduce that the failure area is represented by:

$$\left\{\left(\mathbb{S}\left(X_1\right), \mathbb{S}\left(X_2\right)\right) \in \mathbb{R}_+^2 \mid \bar{\mathbf{C}}\left(\mathbf{F}_1\left(\mathbb{S}\left(X_1\right)\right), \mathbf{F}_2\left(\mathbb{S}\left(X_2\right)\right)\right) \leq \frac{n}{\mathcal{T}}\right\}$$

- We have:

$$\mathcal{T} = \frac{n}{\bar{\mathbf{C}}\left(\mathbf{F}_1\left(\mathbb{S}\left(X_1\right)\right), \mathbf{F}_2\left(\mathbb{S}\left(X_2\right)\right)\right)}$$

  and:

$$\max\left(\mathcal{T}_1, \mathcal{T}_2\right) \leq \mathcal{T} \leq n\mathcal{T}_1\mathcal{T}_2$$

Bivariate stress scenarios

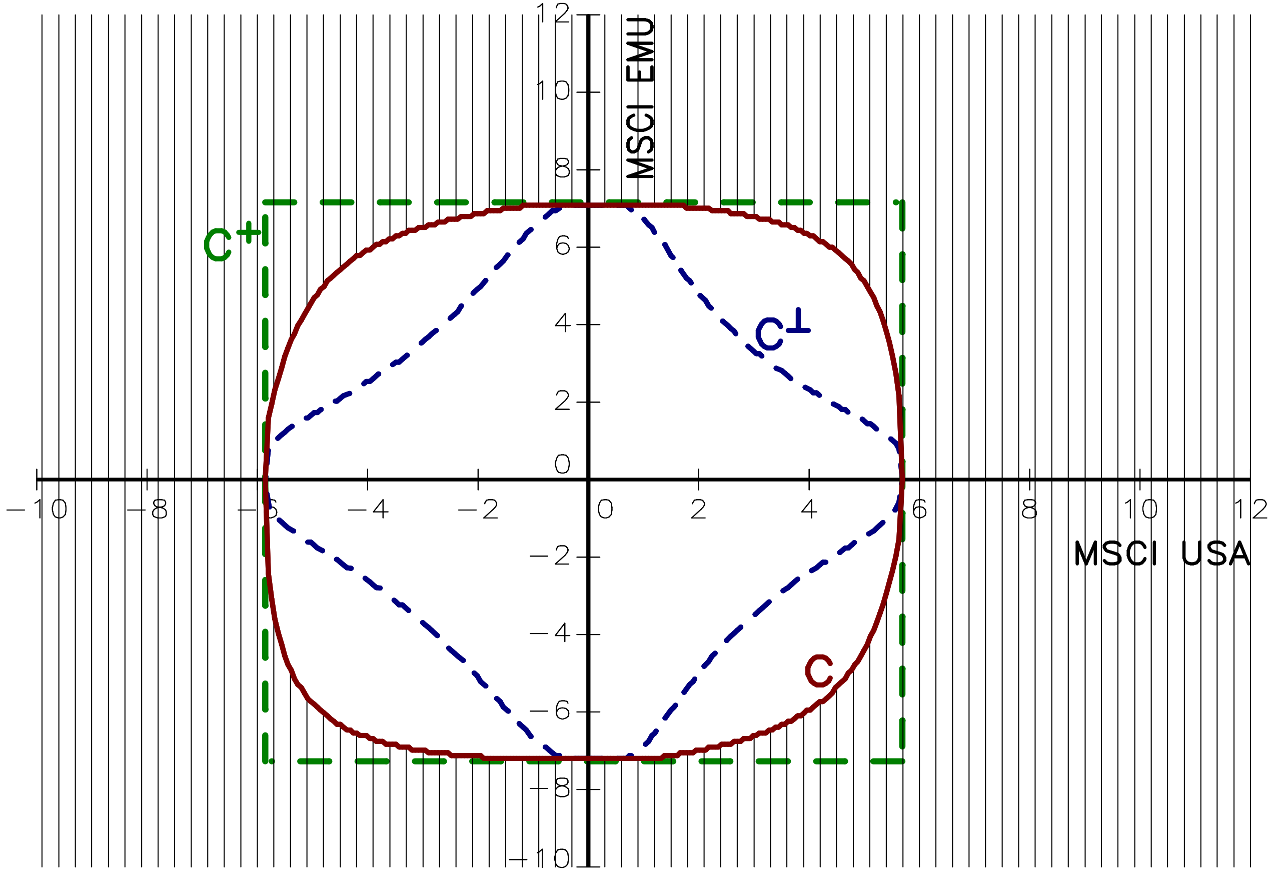


Figure: Failure area of MSCI USA and MSCI EMU indices (blockwise dependence)

Bivariate stress scenarios

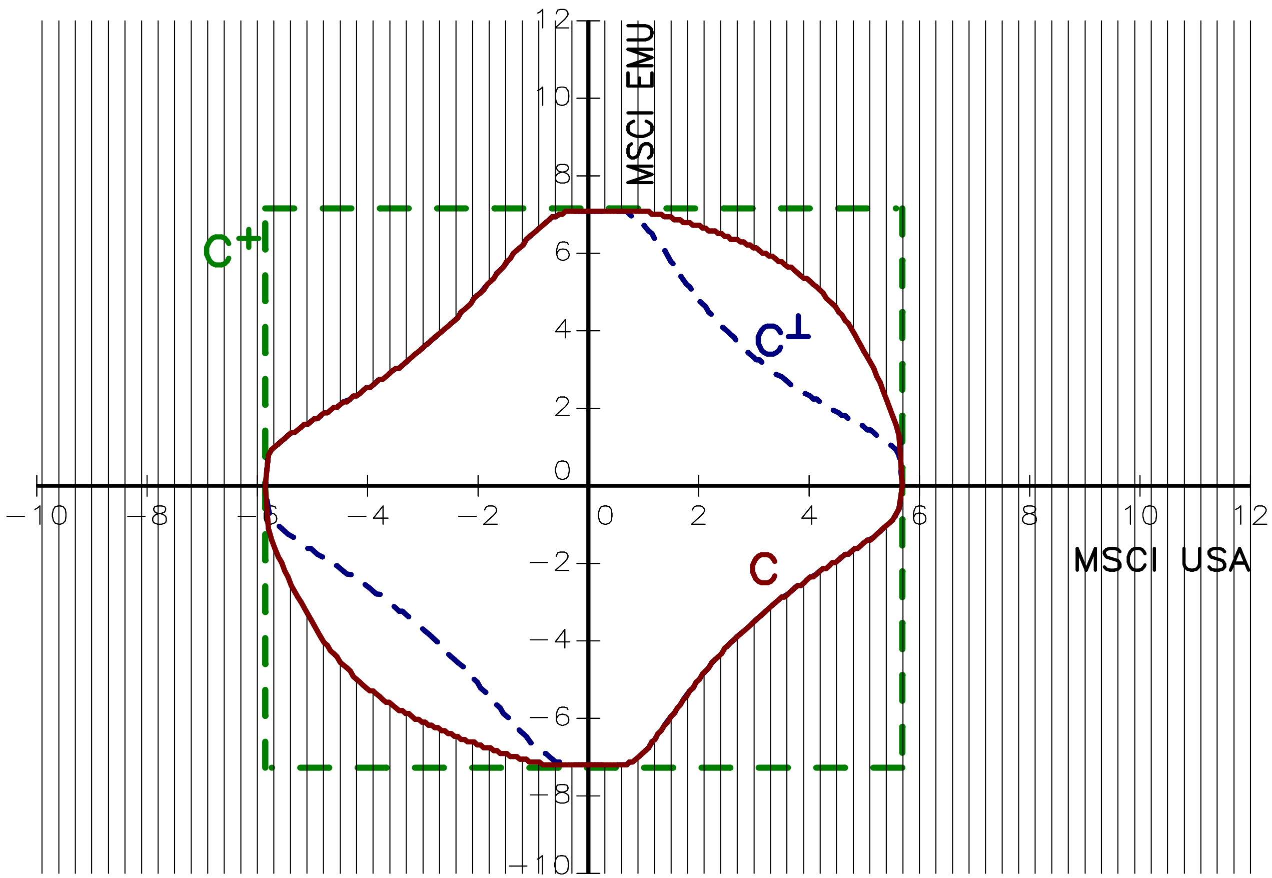


Figure: Failure area of MSCI USA and MSCI EMU indices (daily dependence)

Multivariate stress scenarios

$\Rightarrow$ $\bar{\mathbf{C}}$ has a compliacted expression (see HFRM, Section 14.2.2.2, page 908)

The conditional expectation solution

Given a joint stress scenario $\mathbb{S}(X) = (\mathbb{S}(X_1), \ldots, \mathbb{S}(X_n))$, the conditional stress scenario of Y is:

$$\begin{aligned}
\mathbb{S}(Y) &= \mathbb{E}[Y_t \mid X_t = (\mathbb{S}(X_1), \ldots, \mathbb{S}(X_n))] \\
&= \beta_0 + \sum_{i=1}^{n} \beta_i \mathbb{S}(X_i)
\end{aligned}$$

The conditional expectation solution

Logit transformation

- We use the following transformation:

$$Z_t = \ln \left(\frac{Y_t}{1 - Y_t} \right)$$

- We have:

$$Y_t = \frac{\exp(Z_t)}{1 + \exp(Z_t)} = \frac{1}{1 + \exp(-Z_t)} = h(Z_t)$$

where $h(z)$ is the logit transformation

- We deduce that:

$$\mathbb{E}[Y_t \mid X_t = (x_1, \ldots, x_n)] = \int_{-\infty}^{\infty} h \left(\beta_0 + \sum_{i=1}^{n} \beta_i X_{i,t} + \omega \right) \frac{1}{\sigma} \phi \left(\frac{\omega}{\sigma} \right) \mathrm{d}\omega$$

The conditional expectation solution

Example

- We assume that the probability of default PD_t at time t is explained by the following linear regression model:

$$\ln\left(\frac{\mathrm{PD}_t}{1-\mathrm{PD}_t}\right) = -2.5 - 5g_t - 3\pi_t + 2u_t + \varepsilon_t$$

  where $\varepsilon_t \sim \mathcal{N}(0, 0.25)$, g_t is the growth rate of the GDP, π_t is the inflation rate, and u_t is the unemployment rate
- The baseline scenario is defined by $g_t = 2\%$, $\pi_t = 2\%$ and $u_t = 5\%$
- The stress scenario is equal to $g_t = -8\%$, $\pi_t = 5\%$ and $u_t = 10\%$

The conditional expectation solution

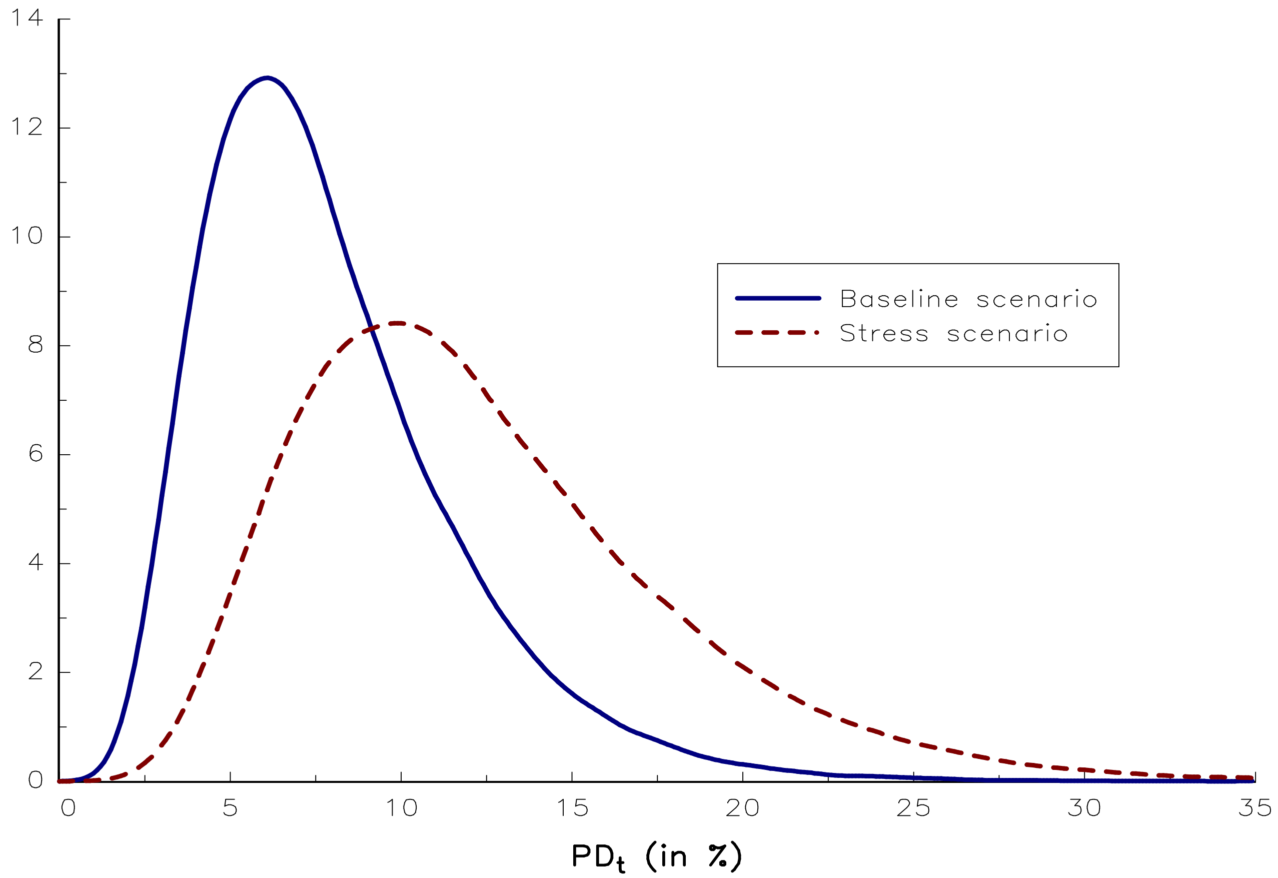


Figure: Probability density function of PD_t

The conditional expectation solution

⇒ The conditional expectation is equal to 7.90% for the baseline scenario and 12.36% for the stress scenario

⇒ The figure of 7.90% can be interpreted as the long-run (or unconditional) probability of default that is used in the IRB formula (i.e. Pillar I)

⇒ The figure of 12.36% may be used in Pillar II

The conditional expectation solution

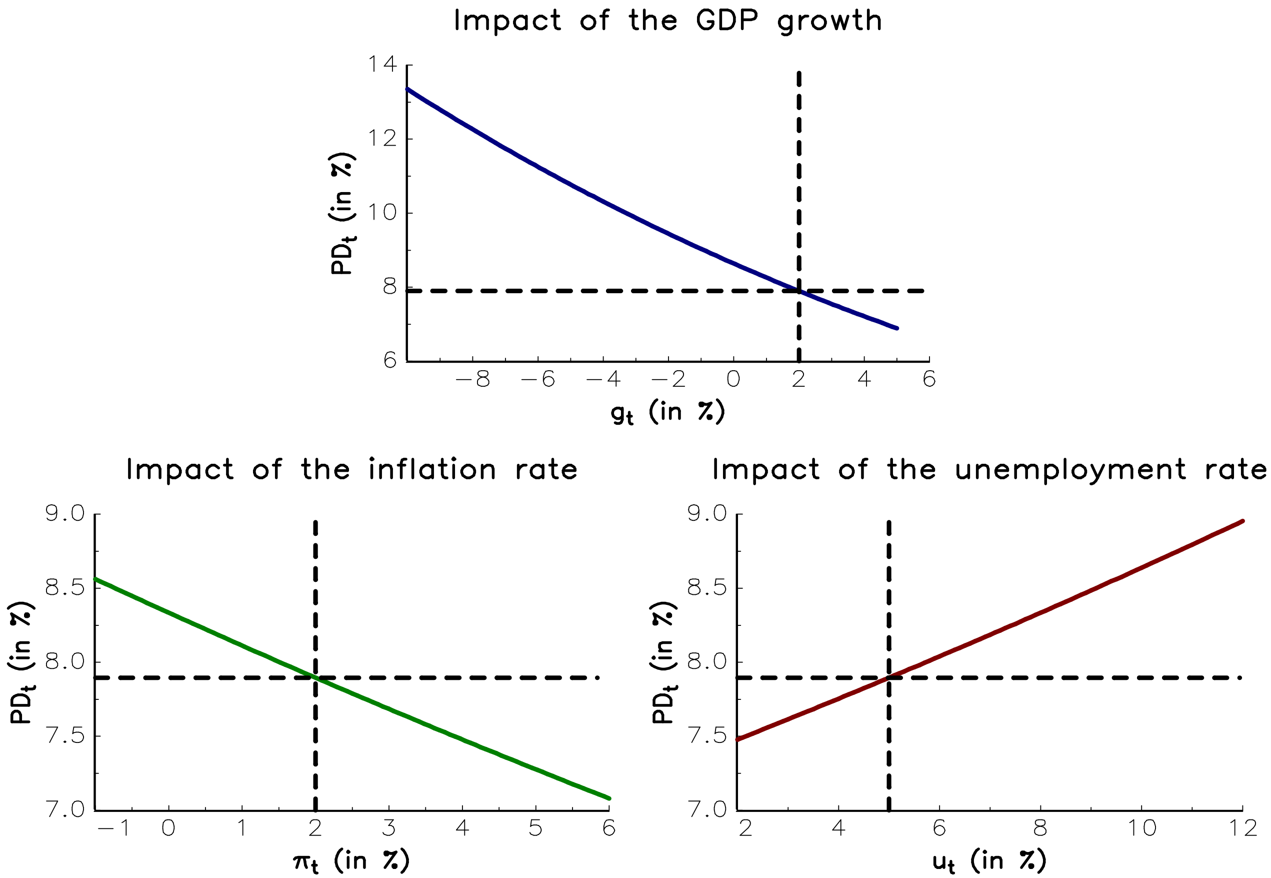


Figure: Relationship between the macroeconomic variables and PD_t

The conditional expectation solution

Table: Stress scenario of the probability of default

t	g_t	π_t	u_t	$\mathbb{E}\left[\mathrm{PD}_t \mid \mathbb{S}(X)\right]$	$q_{90\%}\left(\mathbb{S}(X)\right)$
0	2.00	2.00	5.00	7.90	12.78
1	−6.00	2.00	6.00	11.45	18.26
2	−7.00	1.00	7.00	12.47	19.79
3	−9.00	1.00	9.00	14.03	22.14
4	−7.00	1.00	10.00	13.12	20.78
5	−7.00	2.00	11.00	13.01	20.59
6	−6.00	2.00	10.00	12.26	19.49
7	−4.00	4.00	9.00	10.49	16.80
8	−2.00	3.00	8.00	9.70	15.58
9	−1.00	3.00	7.00	9.11	14.68
10	2.00	3.00	6.00	7.82	12.68
11	4.00	3.00	6.00	7.14	11.60
12	4.00	3.00	6.00	7.14	11.60

The conditional quantile solution

We could also define the conditional stress scenario $\mathbb{S}(Y) = q_\alpha(\mathbb{S}(X))$ as the solution of the quantile regression:

$$\Pr\{Y_t \leq q_\alpha(\mathbb{S}) \mid X_t = \mathbb{S}\} = \alpha$$

The solution is given by:

$$\begin{aligned} \mathbb{S}(Y) &= q_\alpha(\mathbb{S}) \\ &= \mathbf{F}_y^{-1}\left(\mathbf{C}_{2|1}^{-1}(\mathbf{F}_x(\mathbb{S}(X)), \alpha)\right) \end{aligned}$$

⇒ See HFRM, Section 14.2.3.2, pages 912-915

Reverse stress testing

Reverse stress test "*means an institution stress test that starts from the identification of the pre-defined outcome (e.g. points at which an institution business model becomes unviable, or at which the institution can be considered as failing or likely to fail) and then explores scenarios and circumstances that might cause this to occur*"

- In stress testing, extreme scenarios of risk factors are used to test the viability of the bank:

$$\left(\mathbb{S}\left(\mathcal{F}_1\right), \ldots, \mathbb{S}\left(\mathcal{F}_m\right)\right) \Rightarrow \mathbb{S}\left(L\left(w\right)\right) \Rightarrow \begin{cases} D = 0 & \text{if } \mathbb{S}\left(L\left(w\right)\right) < C \\ D = 1 & \text{otherwise} \end{cases}$$

- In reverse stress testing, extreme scenarios of risk factors are deduced from the bankruptcy scenario:

$$D = 1 \Rightarrow \mathbb{RS}\left(L\left(w\right)\right) \Rightarrow \left(\mathbb{RS}\left(\mathcal{F}_1\right), \ldots, \mathbb{RS}\left(\mathcal{F}_m\right)\right)$$

Reverse stress testing

We recall that:

$$L(w) = \ell(\mathcal{F}_1, \ldots, \mathcal{F}_m; w)$$

The reverse stress scenario $\mathbb{RS}$ is the set of risk factors that corresponds to the stressed loss $\mathbb{RS}(L(w))$:

$$\mathbb{RS} = \{(\mathbb{RS}(\mathcal{F}_1), \ldots, \mathbb{RS}(\mathcal{F}_m)) : \ell(\mathbb{S}(\mathcal{F}_1), \ldots, \mathbb{S}(\mathcal{F}_m); w) = \mathbb{RS}(L(w))\}$$

$\Rightarrow$ Not a unique solution

Mathematical solution

We can use the following optimization program

$$\begin{aligned}(\mathbb{RS}(\mathcal{F}_1), \ldots, \mathbb{RS}(\mathcal{F}_m)) &= \arg\max \ln f(\mathcal{F}_1, \ldots, \mathcal{F}_m) \\ &\text{s.t.} \quad \ell(\mathbb{S}(\mathcal{F}_1), \ldots, \mathbb{S}(\mathcal{F}_m); w) = \mathbb{RS}(L(w))\end{aligned}$$

where $f(x_1, \ldots, x_m)$ is the probability density function of the risk factors $(\mathcal{F}_1, \ldots, \mathcal{F}_m)$

Reverse stress testing

We assume that $\mathcal{F} \sim \mathcal{N}(\mu_{\mathcal{F}}, \Sigma_{\mathcal{F}})$ and $L(w) = \sum_{j=1}^{m} w_j \mathcal{F}_j = w^\top \mathcal{F}$. The optimization problem becomes:

$$\begin{aligned} \mathbb{RS}(\mathcal{F}) &= \arg\min \frac{1}{2}(\mathcal{F} - \mu_{\mathcal{F}})^\top \Sigma_{\mathcal{F}}^{-1} (\mathcal{F} - \mu_{\mathcal{F}}) \\ \text{s.t.} &\quad w^\top \mathcal{F} = \mathbb{RS}(L(w)) \end{aligned}$$

The Lagrange function is:

$$\mathcal{L}(\mathcal{F}; \lambda) = \frac{1}{2}(\mathcal{F} - \mu_{\mathcal{F}})^\top \Sigma_{\mathcal{F}}^{-1} (\mathcal{F} - \mu_{\mathcal{F}}) - \lambda \left(w^\top \mathcal{F} - \mathbb{RS}(L(w)) \right)$$

The first-order condition is $\Sigma_{\mathcal{F}}^{-1}(\mathcal{F} - \mu_{\mathcal{F}}) - \lambda w = \mathbf{0}$. It follows that $\mathcal{F} = \mu_{\mathcal{F}} + \lambda \Sigma_{\mathcal{F}} w$, $w^\top \mathcal{F} = w^\top \mu_{\mathcal{F}} + \lambda w^\top \Sigma_{\mathcal{F}} w$, $\lambda = \left(\mathbb{RS}(L(w)) - w^\top \mu_{\mathcal{F}} \right) / w^\top \Sigma_{\mathcal{F}} w$ and:

$$\mathbb{RS}(\mathcal{F}) = \mu_{\mathcal{F}} + \frac{\Sigma_{\mathcal{F}} w}{w^\top \Sigma_{\mathcal{F}} w} \left(\mathbb{RS}(L(w)) - w^\top \mu_{\mathcal{F}} \right)$$

Reverse stress testing

Another approach for solving the inverse problem is to consider the joint distribution of $\mathcal{F}$ and $L(w)$:

$$\begin{pmatrix} \mathcal{F} \\ L(w) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_{\mathcal{F}} \\ w^\top \mu_{\mathcal{F}} \end{pmatrix}, \begin{pmatrix} \Sigma_{\mathcal{F}} & \Sigma_{\mathcal{F}} w \\ w^\top \Sigma_{\mathcal{F}} & w^\top \Sigma_{\mathcal{F}} w \end{pmatrix}\right)$$

The conditional distribution of $\mathcal{F}$ given $L(w) = \mathbb{RS}(L(w))$ is Gaussian:

$$\mathcal{F} \mid L(w) = \mathbb{RS}(L(w)) \sim \mathcal{N}\left(\mu_{\mathcal{F}\mid L(w)}, \Sigma_{\mathcal{F}\mid L(w)}\right)$$

We know that the maximum of the probability density function of the multivariate normal distribution is reached when the random vector is exactly equal to the mean. We deduce that:

$$\mathbb{RS}(\mathcal{F}) = \mu_{\mathcal{F}\mid L(w)} = \mu_{\mathcal{F}} + \frac{\Sigma_{\mathcal{F}} w}{w^\top \Sigma_{\mathcal{F}} w}\left(\mathbb{RS}(L(w)) - w^\top \mu_{\mathcal{F}}\right)$$

Reverse stress testing

Example

We assume that $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2)$, $\mu_{\mathcal{F}} = (5, 8)$, $\sigma_{\mathcal{F}} = (1.5, 3.0)$ and $\rho(\mathcal{F}_1, \mathcal{F}_2) = -50\%$. The sensitivity vector w to the risk factors is equal to $(10, 3)$

The stress scenario is the collection of univariate stress scenarios at the 99% confidence level:

$$\begin{aligned} \mathbb{S}(\mathcal{F}_1) &= 5 + 1.5 \cdot \Phi^{-1}(99\%) = 8.49 \\ \mathbb{S}(\mathcal{F}_2) &= 8 + 3.0 \cdot \Phi^{-1}(99\%) = 14.98 \end{aligned}$$

The stressed loss is then equal to:

$$\mathbb{S}(L(w)) = 10 \cdot 8.49 + 3 \cdot 14.98 = 129.53$$

Reverse stress testing

We assume that the reverse stressed loss is equal to 129.53 $\Rightarrow$ we deduce that $\mathbb{RS}(\mathcal{F}_1) = 10.14$ and $\mathbb{RS}(\mathcal{F}_2) = 9.47$

Remark

The reverse stress scenario is very different than the stress scenario even if they give the same loss. In fact, we have $f(\mathbb{S}(\mathcal{F}_1), \mathbb{S}(\mathcal{F}_2)) = 0.8135 \cdot 10^{-6}$ *and* $f(\mathbb{RS}(\mathcal{F}_1), \mathbb{RS}(\mathcal{F}_2)) = 4.4935 \cdot 10^{-6}$*, meaning that the occurrence probability of the reverse stress scenario is more than five times higher than the occurrence probability of the stress scenario*

Reverse stress testing

In the general case, we consider the following optimization problem:

$$\begin{aligned} \left(\mathbb{RS}\left(\mathcal{F}_1\right), \ldots, \mathbb{RS}\left(\mathcal{F}_m\right)\right) &= \arg\max \ln f\left(\mathcal{F}_1, \ldots, \mathcal{F}_m\right) \\ &\text{s.t.} \quad \ell\left(\mathbb{S}\left(\mathcal{F}_1\right), \ldots, \mathbb{S}\left(\mathcal{F}_m\right); w\right) \geq \mathbb{RS}\left(L\left(w\right)\right) \end{aligned}$$

and we use the Monte Carlo simulation method to estimate the reverse stress scenario

Hard to implement in practice!

Exercises

- Exercise 14.3.1 – Construction of a stress scenario with the GEV distribution

References

Basel Committee on Banking Supervision (2017)
Supervisory and Bank Stress Testing: Range of Practices, December 2017.

Roncalli, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 14.

Roncalli, T. (2020)
Handbook of Financial Risk Management – Companion Book, Chapter 14.

Course 2023-2024 in Financial Risk Management
Lecture 12. Credit Scoring Models

Thierry Roncalli*

*Amundi Asset Management[23]

*University of Paris-Saclay

September 2023

[23] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

Credit scoring

- Credit scoring refers to statistical models to measure the creditworthiness of a person or a company
- Mortgage, credit card, personal loan, etc.
- Credit scoring first emerged in the United States
- The FICO score was introduced in 1989 by Fair Isaac Corporation

Judgmental credit systems versus credit scoring systems

- In 1941, Durand presented a statistical analysis of credit valuation
- He showed that credit analysts uses similar factors, and proposed a credit rating formula based on nine factors: (1) age, (2) sex, (3) stability of residence, (4) occupation, (5) industry, (6) stability of employment, (7) bank account, (8) real estate and (9) life insurance
- The score is additive and can take values between 0 and 3.46
- From an industrial point of view, a credit scoring system has two main advantages compared to a judgmental credit system:
  1. it is cost efficient, and can treat a huge number of applicants;
  2. decision-making process is rapid and consistent across customers.

Scoring models for corporate bankruptcy

Altman Z score model (1968)

- The score was equal to:

$$Z = 1.2 \cdot X_1 + 1.4 \cdot X_2 + 3.3 \cdot X_3 + 0.6 \cdot X_4 + 1.0 \cdot X_5$$

- The variables X_j represent the following financial ratios:

X_j	Ratio
X_1	Working capital / Total assets
X_2	Retained earnings / Total assets
X_3	Earnings before interest and tax / Total assets
X_4	Market value of equity / Total liabilities
X_5	Sales / Total assets

- If we note Z_i the score of the firm i, we can calculate the normalized score:

$$Z_i^\star = (Z_i - m_z) / \sigma_z$$

where m_z and σ_z are the mean and standard deviation of the observed scores

- A low value of $Z_i^\star$ (for instance $Z_i^\star < 2.5$) indicates that the firm has a high probability of default

New developments

- Default of corporate firms
- Consumer credit and retail debt management (credit cards, mortgages, etc.)
- Statistical methods: discriminant analysis, logistic regression, survival model, machine learning techniques

Choice of the risk factors

The five Cs:

1. **Capacity** measures the applicant's ability to meet the loan payments (e.g., debt-to-income, job stability, cash flow dynamics)
2. **Capital** is the size of assets that are held by the borrower (e.g. net wealth of the borrower)
3. **Character** measures the willingness to repay the loan (e.g. payment history of the applicant)
4. **Collateral** concerns additional forms of security that the borrower can provide to the lender
5. **Conditions** refer to the characteristics of the loan and the economic conditions that might affect the borrower (e.g. maturity, interests paid)

Choice of the risk factors

Table: An example of risk factors for consumer credit

Character	Age of applicant
	Marital status
	Number of children
	Educational background
	Time with bank
	Time at present address
Capacity	Annual income
	Current living expenses
	Current debts
	Time with employer
Capital	Purpose of the loan
	Home status
	Saving account
Condition	Maturity of the loan
	Paid interests

Choice of the risk factors

- Scores are developed by banks and financial institutions, but they can also be developed by consultancy companies
- This is the case of the FICO® scores, which are the most widely used credit scoring systems in the world

5 main categories

1. Payment history (35%)
2. Amount of debt (30%)
3. Length of credit history (15%)
4. New credit (10%)
5. Credit mix (10%)

Range

Generally from 300 to 850 (average score of US consumers is 695)

- Exceptional (800+)
- Very good (740-799)
- Good (670-739)
- Fair (580-669)
- Poor (580−)

Choice of the risk factors

Corporate credit scoring systems use financial ratios:

1. **Profitability**: gross profit margin, operating profit margin, return-on-equity (ROE), etc.
2. **Solvency**: debt-to-assets ratio, debt-to-equity ratio, interest coverage ratio, etc.
3. **Leverage**: liabilities-to-assets ratio (financial leverage ratio), long-term debt/assets, etc.
4. **Liquidity**: current assets/current liabilities (current ratio), quick assets/current liabilities (quick or cash ratio), total net working capital, assets with maturities of less than one year, etc.

Data preparation

- Check the data and remove outliers or fill missing values
- Variable transformation
- Slicing-and-dicing segmentation
- Potential interaction

Variable selection

- Many candidate variables $X = (X_1, \ldots, X_m)$ for explaining the variable Y
- The variable selection problem consists in finding the best set of optimal variables
- We assume the following statistical model:

$$Y = f(X) + u$$

where $u \sim \mathcal{N}(0, \sigma^2)$

- We denote the prediction by $\hat{Y} = \hat{f}(X)$. We have:

$$\begin{aligned}
\mathbb{E}\left[\left(Y - \hat{Y}\right)^2\right] &= \mathbb{E}\left[\left(f(X) + u - \hat{f}(X)\right)^2\right] \\
&= \left(\mathbb{E}\left[\hat{f}(X)\right] - f(X)\right)^2 + \mathbb{E}\left[\left(\hat{f}(X) - \mathbb{E}\left[\hat{f}(X)\right]\right)^2\right] + \sigma^2 \\
&= \mathrm{Bias}^2 + \mathrm{Variance} + \mathrm{Error}
\end{aligned}$$

Variable selection

- Best subset selection:

$$\text{AIC}\left(\alpha\right) = -2\ell_{(k)}\left(\hat{\theta}\right) + \alpha \cdot \text{df}_{(k)}^{(\text{model})}$$

- Stepwise approach:

$$F = \frac{\text{RSS}\left(\hat{\theta}_{(k)}\right) - \text{RSS}\left(\hat{\theta}_{(k+1)}\right)}{\text{RSS}\left(\hat{\theta}_{(k+1)}\right) / \text{df}_{(k+1)}^{(\text{residual})}}$$

- Lasso approach:

$$y_i = \sum_{k=1}^{K} \beta_k x_{i,k} + u_i \quad \text{s.t.} \quad \sum_{k=1}^{K} \left|\beta_k\right| \leq \tau$$

Score modeling, validation and follow-up

- Cross-validation approach (leave-p-out cross-validation or LpOCV, leave-one-out cross-validation or LOOCV, Press statistic)
- Score modeling
  - $S = f\left(X; \hat{\theta}\right)$ is the score
  - Decision rule:

$$\left\{ \begin{array}{l} S < s \Longrightarrow Y = 0 \Longrightarrow \text{reject} \\ S \geq s \Longrightarrow Y = 1 \Longrightarrow \text{accept} \end{array} \right.$$

- Score follow-up
  - Stability
  - Rejected applicants (reject inference)
  - Backtesting

Statistical methods

- Unsupervised learning is a branch of statistical learning, where test data does not include a response variable
- It is opposed to supervised learning, whose goal is to predict the value of the response variable Y given a set of explanatory variables X
- In the case of unsupervised learning, we only know the X-values, because the Y-values do not exist or are not observed
- Supervised and unsupervised learning are also called '*learning with/without a teacher*' (Hastie *et al.*, 2009)

Clustering

- K-means clustering
- Hierarchical clustering

Clustering

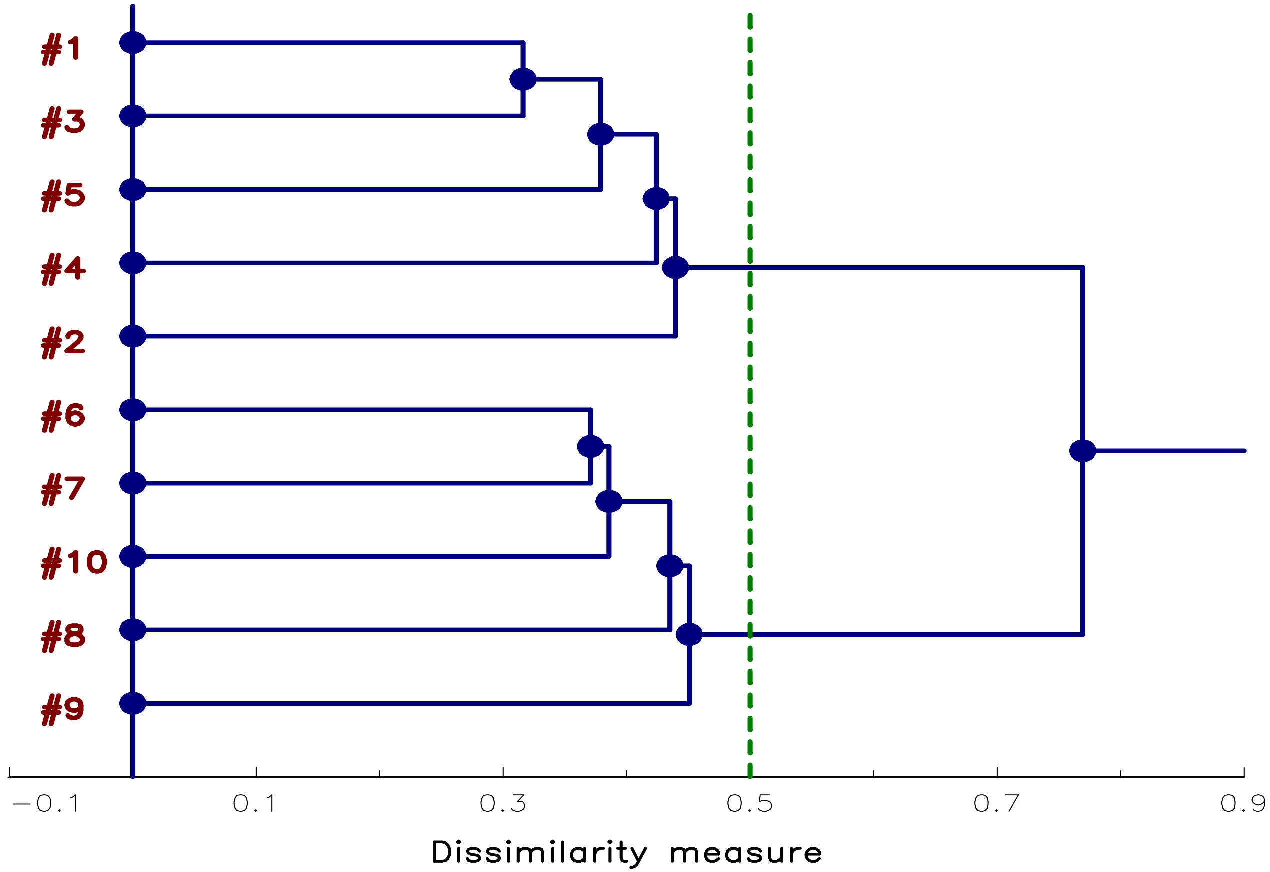


Figure: An example of dendrogram

Dimension reduction

- Principal component analysis
- Non-negative matrix factorization

Discriminant analysis

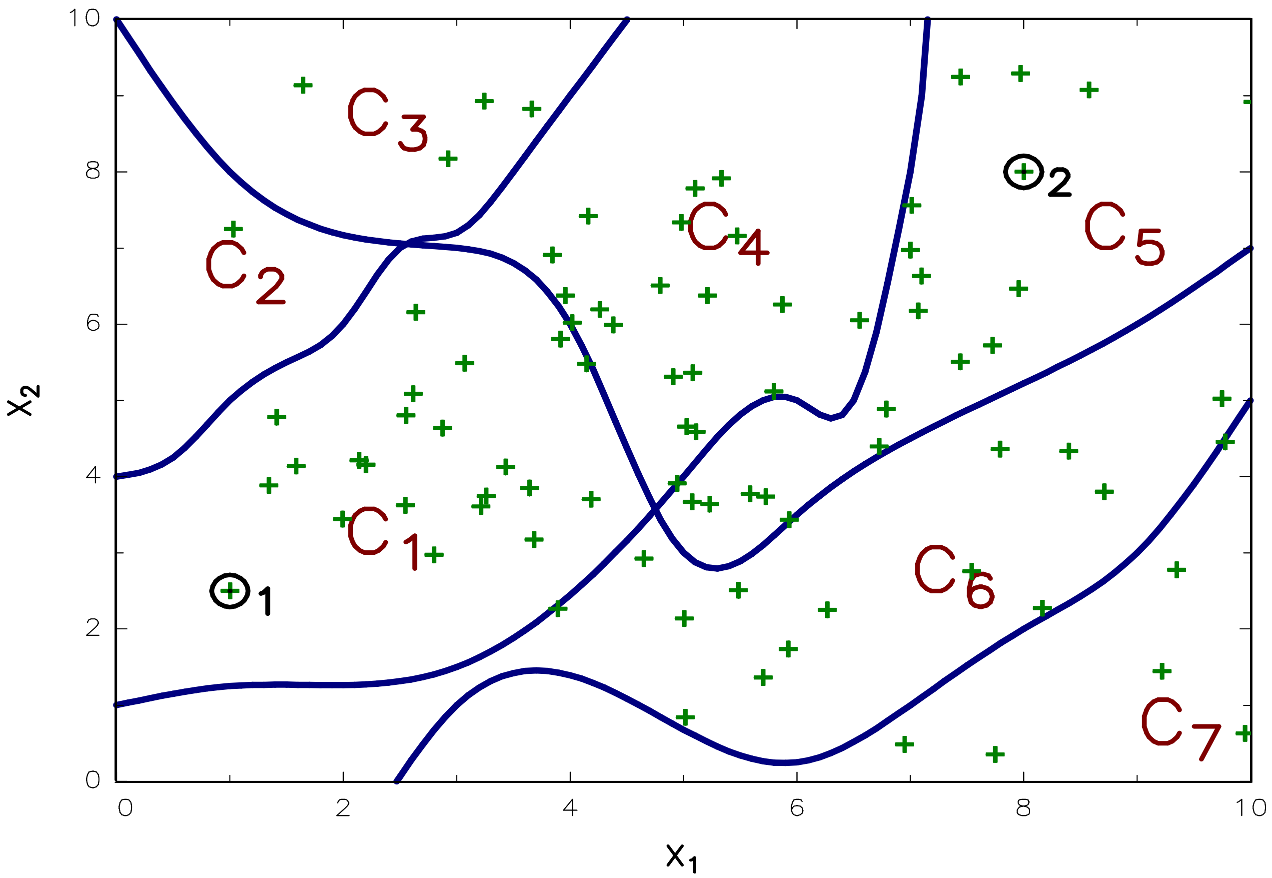


Figure: Classification statistical problem

Discriminant analysis

The two-dimensional case

- Using the Bayes theorem, we have:

$$\Pr\{A \cap B\} = \Pr\{A \mid B\} \cdot \Pr\{B\} = \Pr\{B \mid A\} \cdot \Pr\{A\}$$

- It follows that:

$$\Pr\{A \mid B\} = \Pr\{B \mid A\} \cdot \frac{\Pr\{A\}}{\Pr\{B\}}$$

- If we apply this result to the conditional probability $\Pr\{i \in \mathcal{C}_1 \mid X = x\}$, we obtain:

$$\Pr\{i \in \mathcal{C}_1 \mid X = x\} = \Pr\{X = x \mid i \in \mathcal{C}_1\} \cdot \frac{\Pr\{i \in \mathcal{C}_1\}}{\Pr\{X = x\}}$$

Discriminant analysis
The two-dimensional case

- The log-probability ratio is then equal to:

$$\begin{aligned}\ln \frac{\Pr\{i \in \mathcal{C}_1 \mid X = x\}}{\Pr\{i \in \mathcal{C}_2 \mid X = x\}} &= \ln \left(\frac{\Pr\{X = x \mid i \in \mathcal{C}_1\}}{\Pr\{X = x \mid i \in \mathcal{C}_2\}} \cdot \frac{\Pr\{i \in \mathcal{C}_1\}}{\Pr\{i \in \mathcal{C}_2\}} \right) \\ &= \ln \frac{f_1(x)}{f_2(x)} + \ln \frac{\pi_1}{\pi_2}\end{aligned}$$

  where $\pi_j = \Pr\{i \in \mathcal{C}_j\}$ is the probability of the j^{th} class and $f_j(x) = \Pr\{X = x \mid i \in \mathcal{C}_j\}$ is the conditional pdf of X

- By construction, the decision boundary is defined such that we are indifferent to an assignment rule ($i \in \mathcal{C}_1$ and $i \in \mathcal{C}_2$), implying that:

$$\Pr\{i \in \mathcal{C}_1 \mid X = x\} = \Pr\{i \in \mathcal{C}_2 \mid X = x\} = \frac{1}{2}$$

- Finally, we deduce that the decision boundary satisfies the following equation:

$$\ln \frac{f_1(x)}{f_2(x)} + \ln \frac{\pi_1}{\pi_2} = 0$$

Discriminant analysis

Quadratic discriminant analysis (QDA)

- If we model each class density as a multivariate normal distribution:

$$X \mid i \in \mathcal{C}_j \sim \mathcal{N}\left(\mu_j, \Sigma_j\right)$$

we have:

$$f_j\left(x\right) = \frac{1}{\left(2\pi\right)^{K/2} \left|\Sigma_j\right|^{1/2}} \exp\left(-\frac{1}{2}\left(x-\mu_j\right)^\top \Sigma_j^{-1}\left(x-\mu_j\right)\right)$$

- We deduce that:

$$\ln \frac{f_1\left(x\right)}{f_2\left(x\right)} = \frac{1}{2}\ln\frac{\left|\Sigma_2\right|}{\left|\Sigma_1\right|} - \frac{1}{2}\left(x-\mu_1\right)^\top \Sigma_1^{-1}\left(x-\mu_1\right) + \frac{1}{2}\left(x-\mu_2\right)^\top \Sigma_2^{-1}\left(x-\mu_2\right)$$

- The decision boundary is then given by:

$$\frac{1}{2}\ln\frac{\left|\Sigma_2\right|}{\left|\Sigma_1\right|} - \frac{1}{2}\left(x-\mu_1\right)^\top \Sigma_1^{-1}\left(x-\mu_1\right) + \frac{1}{2}\left(x-\mu_2\right)^\top \Sigma_2^{-1}\left(x-\mu_2\right) + \ln\frac{\pi_1}{\pi_2} = 0$$

Discriminant analysis

Linear discriminant analysis (LDA)

- If we assume that $\Sigma_1 = \Sigma_2 = \Sigma$, we obtain:

$$\frac{1}{2}(x-\mu_2)^\top \Sigma^{-1}(x-\mu_2) - \frac{1}{2}(x-\mu_1)^\top \Sigma^{-1}(x-\mu_1) + \ln\frac{\pi_1}{\pi_2} = 0$$

- We deduce that:

$$(\mu_2-\mu_1)^\top \Sigma^{-1} x = \frac{1}{2}\left(\mu_2^\top \Sigma^{-1}\mu_2 - \mu_1^\top \Sigma^{-1}\mu_1\right) + \ln\frac{\pi_2}{\pi_1}$$

- The decision boundary is then linear in x (and not quadratic)

Discriminant analysis

Example #1

We consider two classes and two explanatory variables $X = (X_1, X_2)$ where $\pi_1 = 50\%$, $\pi_2 = 1 - \pi_1 = 50\%$, $\mu_1 = (1, 3)$, $\mu_2 = (4, 1)$, $\Sigma_1 = I_2$ and $\Sigma_2 = \gamma I_2$ where $\gamma = 1.5$.

Discriminant analysis

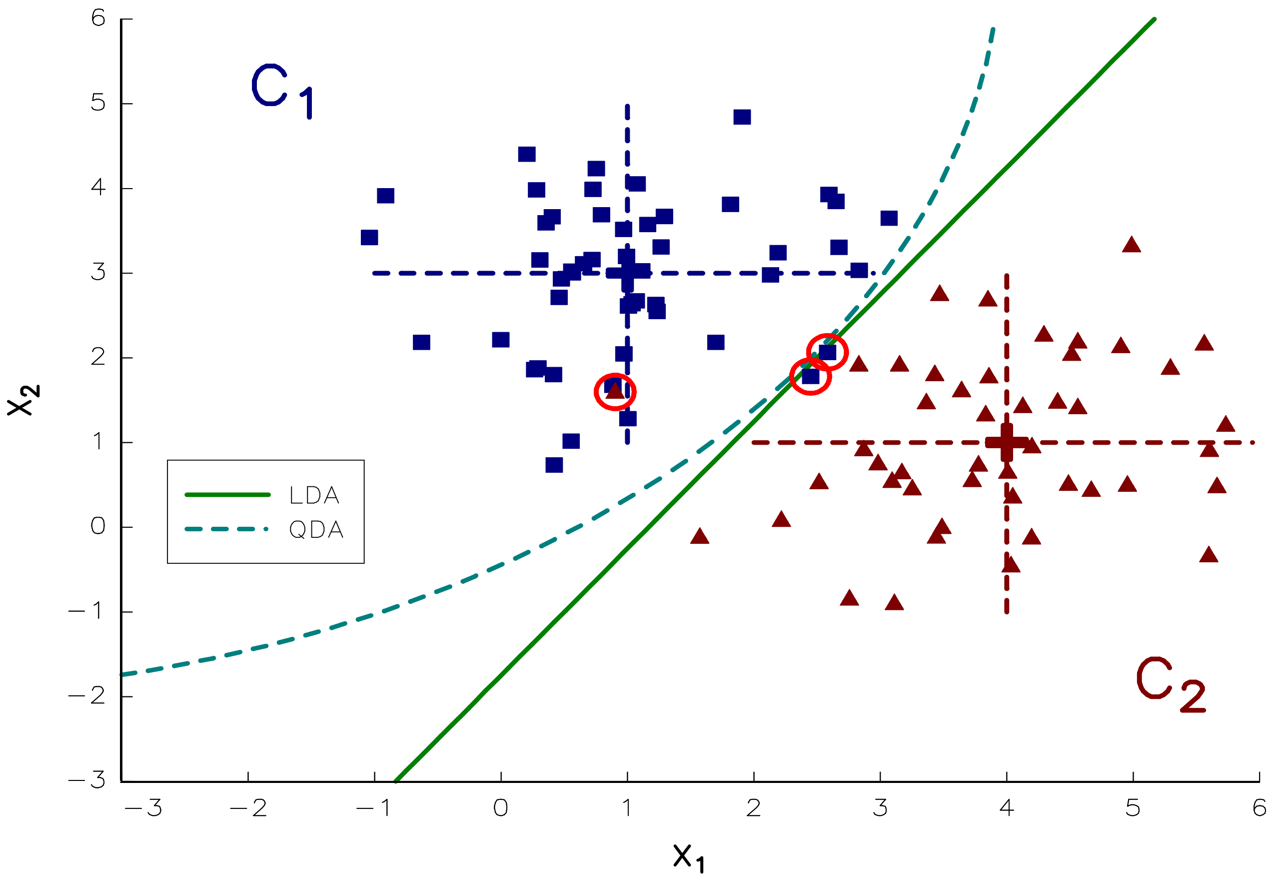


Figure: Boundary decision of discriminant analysis

Discriminant analysis

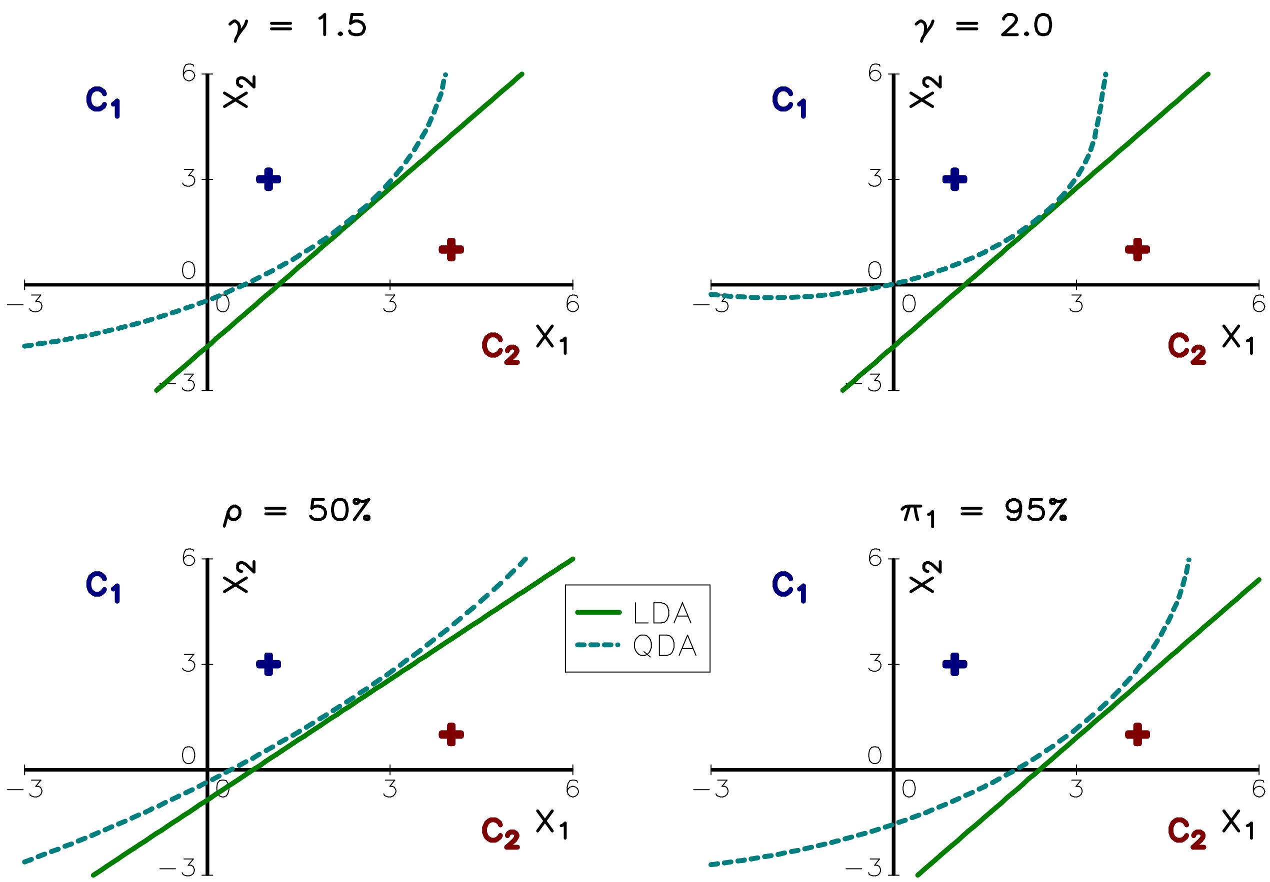


Figure: Impact of the parameters on LDA/QDA boundary decisions

Discriminant analysis
The general case

- We can generalize the previous analysis to J classes
- The Bayes formula gives:

$$\begin{aligned}\Pr\left\{i \in \mathcal{C}_j \mid X = x\right\} &= \Pr\left\{X = x \mid i \in \mathcal{C}_j\right\} \cdot \frac{\Pr\left\{i \in \mathcal{C}_j\right\}}{\Pr\left\{X = x\right\}} \\ &= c \cdot f_j\left(x\right) \cdot \pi_j\end{aligned}$$

  where $c = 1/\Pr\left\{X = x\right\}$ is a normalization constant that does not depend on j
- We note $S_j\left(x\right) = \ln \Pr\left\{i \in \mathcal{C}_j \mid X = x\right\}$ the discriminant score function for the j^{th} class
- We have:

$$S_j\left(x\right) = \ln c + \ln f_j\left(x\right) + \ln \pi_j$$

Discriminant analysis
The general case

- If we again assume that $X \mid i \in \mathcal{C}_j \sim \mathcal{N}\left(\mu_j, \Sigma_j\right)$, the QDA score function is:

$$
\begin{aligned}
S_j\left(x\right) &= \ln c' + \ln \pi_j - \frac{1}{2}\ln\left|\Sigma_j\right| - \frac{1}{2}\left(x-\mu_j\right)^\top \Sigma_j^{-1}\left(x-\mu_j\right) \\
&\propto \ln \pi_j - \frac{1}{2}\ln\left|\Sigma_j\right| - \frac{1}{2}\left(x-\mu_j\right)^\top \Sigma_j^{-1}\left(x-\mu_j\right)
\end{aligned}
$$

  where $\ln c' = \ln c - \frac{K}{2}\ln 2\pi$

- Given an input x, we calculate the scores $S_j\left(x\right)$ for $j = 1,\ldots,J$ and we choose the label $j^\star$ with the highest score value

Discriminant analysis

The general case

- If we assume an homoscedastic model ($\Sigma_j = \Sigma$), the LDA score function becomes:

$$\begin{aligned} S_j(x) &= \ln c'' + \ln \pi_j - \frac{1}{2}(x - \mu_j)^\top \Sigma_j^{-1}(x - \mu_j) \\ &\propto \ln \pi_j + \mu_j^\top \Sigma^{-1} x - \frac{1}{2}\mu_j^\top \Sigma^{-1} \mu_j \end{aligned}$$

where $\ln c'' = \ln c' - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}x^\top \Sigma^{-1} x$

Remark

In practice, the parameters π_j, μ_j and Σ_j are unknown. We replace them by the corresponding estimates $\hat{\pi}_j$, $\hat{\mu}_j$ and $\hat{\Sigma}_j$. For the linear discriminant analysis, $\hat{\Sigma}$ is estimated by pooling all the classes.

Discriminant analysis
The general case

Example #2

We consider the classification problem of 33 observations with two explanatory variables X_1 and X_2, and three classes $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$:

i	$\mathcal{C}_j$	X_1	X_2	i	$\mathcal{C}_j$	X_1	X_2	i	$\mathcal{C}_j$	X_1	X_2
1	1	1.03	2.85	12	2	3.70	5.08	23	3	3.55	0.58
2	1	0.20	3.30	13	2	2.81	1.99	24	3	3.86	1.83
3	1	1.69	3.73	14	2	3.66	2.61	25	3	5.39	0.47
4	1	0.98	3.52	15	2	5.63	4.19	26	3	3.15	−0.18
5	1	0.98	5.15	16	2	3.35	3.64	27	3	4.93	1.91
6	1	3.47	6.56	17	2	2.97	3.55	28	3	3.87	2.61
7	1	3.94	4.68	18	2	3.16	2.92	29	3	4.09	1.43
8	1	1.55	5.99	19	3	3.00	0.98	30	3	3.80	2.11
9	1	1.15	3.60	20	3	3.09	1.99	31	3	2.79	2.10
10	2	1.20	2.27	21	3	5.45	0.60	32	3	4.49	2.71
11	2	3.66	5.49	22	3	3.59	−0.46	33	3	3.51	1.82

Discriminant analysis

The general case

Table: Parameter estimation of the discriminant analysis

Class	$\mathcal{C}_1$		$\mathcal{C}_2$		$\mathcal{C}_3$	
$\hat{\pi}_j$	0.273		0.273		0.455	
$\hat{\mu}_j$	1.666	4.376	3.349	3.527	3.904	1.367
$\hat{\Sigma}_j$	1.525	0.929	1.326	0.752	0.694	−0.031
	0.929	1.663	0.752	1.484	−0.031	0.960

For the LDA method, we have:

$$\hat{\Sigma} = \begin{pmatrix} 1.91355 & -0.71720 \\ -0.71720 & 3.01577 \end{pmatrix}$$

Discriminant analysis

The general case

Table: Computation of the discriminant scores $S_j(x)$

i	QDA			LDA			LDA^2		
	$S_1(x)$	$S_2(x)$	$S_3(x)$	$S_1(x)$	$S_2(x)$	$S_3(x)$	$S_1(x)$	$S_2(x)$	$S_3(x)$
1	−2.28	−3.69	−7.49	0.21	−0.96	−0.79	6.93	5.60	5.76
2	−2.28	−6.36	−12.10	−0.26	−2.17	−2.34	1.38	−2.13	−1.89
3	−1.76	−3.13	−6.79	2.84	2.16	1.71	12.13	12.01	11.38
4	−1.80	−4.43	−8.88	1.35	0.09	−0.22	7.73	6.20	5.93
5	−2.36	−7.75	−13.70	4.32	2.93	1.45	8.12	5.54	4.76
6	−3.16	−5.63	−14.68	10.75	11.36	8.95	14.82	13.99	12.96
7	−3.79	−1.92	−6.32	8.06	9.22	8.15	17.36	19.03	17.89
8	−2.85	−8.43	−15.23	6.73	5.76	3.70	10.47	8.09	7.15
9	−1.74	−4.12	−8.37	1.76	0.64	0.27	8.94	7.77	7.39
10	−3.14	−3.21	−6.17	−0.58	−1.56	−0.98	6.59	5.55	6.15
11	−2.87	−3.01	−9.45	9.10	9.96	8.31	16.89	17.65	16.42
12	−3.04	−2.38	−7.77	8.42	9.34	7.98	17.28	18.50	17.28
13	−6.32	−2.29	−1.62	1.41	1.82	2.64	12.48	13.94	14.46
14	−6.91	−2.07	−1.42	3.86	4.94	5.34	15.15	17.41	17.34
15	−9.79	−3.62	−7.12	9.79	12.43	11.75	12.58	14.01	13.50
16	−3.90	−1.47	−3.44	5.25	5.99	5.65	16.84	18.82	18.03
17	−3.31	−1.55	−3.61	4.50	4.92	4.63	16.25	17.95	17.21
18	−4.84	−1.60	−2.19	3.65	4.28	4.45	15.51	17.48	17.14
19	−10.21	−4.12	−1.27	−0.13	0.52	2.06	8.98	9.99	11.70
20	−7.05	−2.41	−1.24	1.85	2.50	3.32	12.99	14.72	15.22
21	−23.11	−11.16	−2.56	2.98	5.75	7.61	3.79	4.57	7.26
22	−19.22	−9.53	−2.42	−1.84	−0.57	2.01	1.81	1.53	5.51
23	−13.86	−5.92	−1.01	−0.01	1.15	2.98	7.65	8.67	10.95
24	−10.01	−3.43	−0.70	2.75	4.07	5.02	12.84	14.95	15.65
25	−23.48	−11.44	−2.54	2.65	5.38	7.33	3.40	4.09	6.95
26	−15.87	−7.59	−2.30	−2.01	−1.14	1.23	3.19	3.02	6.50
27	−14.09	−5.40	−1.52	4.56	6.78	7.70	11.17	13.24	14.08
28	−7.55	−2.27	−1.39	4.18	5.45	5.85	15.10	17.44	17.40
29	−12.40	−4.67	−0.61	2.38	3.92	5.17	11.21	13.14	14.33
30	−8.85	−2.87	−0.88	3.17	4.41	5.17	13.77	15.97	16.37
31	−5.97	−2.17	−1.72	1.58	1.97	2.70	12.78	14.26	14.67
32	−9.40	−2.97	−1.81	5.33	7.11	7.46	14.55	16.95	16.93
33	−8.84	−3.01	−0.80	2.19	3.21	4.16	12.82	14.77	15.45

Discriminant analysis

The general case

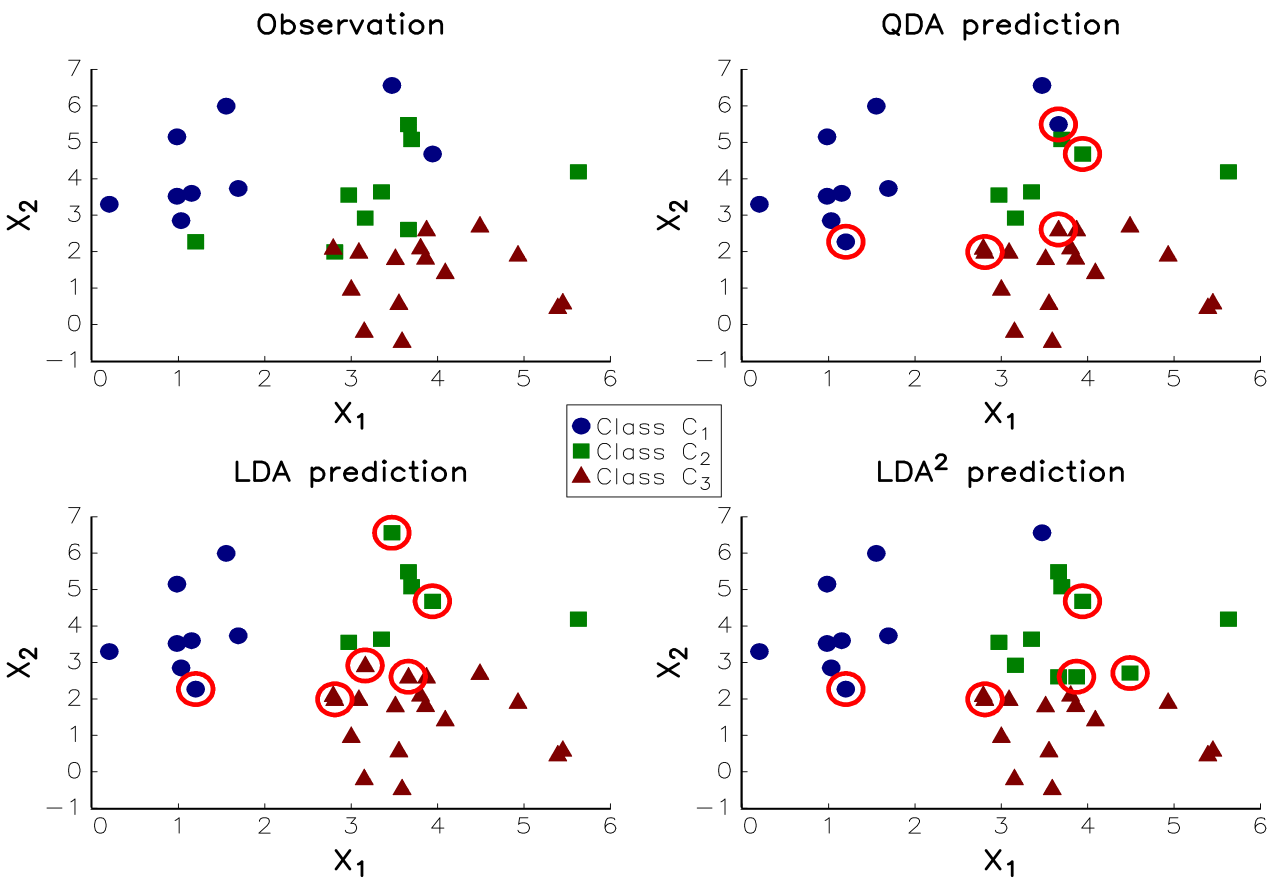


Figure: Comparing QDA, LDA and LDA^2 predictions

Discriminant analysis

The general case

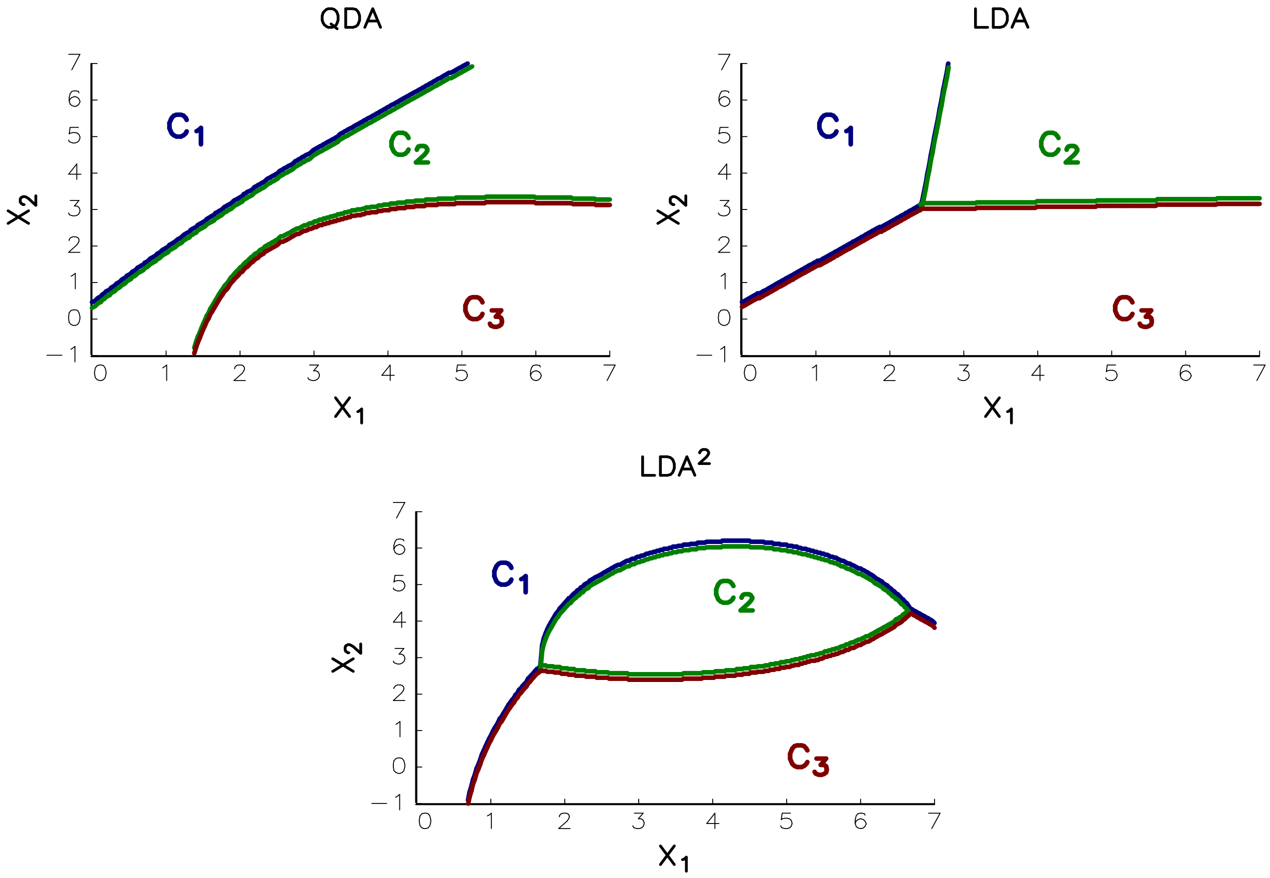


Figure: QDA, LDA and LDA^2 decision regions

Discriminant analysis

Class separation maximization

- We note $x_i = (x_{i,1}, \ldots, x_{i,K})$ the $K \times 1$ vector of exogenous variables X for the i^{th} observation
- The mean vector and the variance (or scatter) matrix of Class $\mathcal{C}_j$ is equal to $\hat{\mu}_j = \frac{1}{n_j} \sum_{i \in \mathcal{C}_j} x_i$ and $\mathbf{S}_j = n\hat{\Sigma}_j = \sum_{i \in \mathcal{C}_j} (x_i - \hat{\mu}_j)(x_i - \hat{\mu}_j)^\top$ where n_j is the number of observations in the j^{th} class
- If consider the total population, we also have $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $\mathbf{S} = n\hat{\Sigma} = \sum_{i=1}^{n} (x_i - \hat{\mu})(x_i - \hat{\mu})^\top$

Discriminant analysis

Class separation maximization

- We notice that:

$$\hat{\mu} = \frac{1}{n} \sum_{j=1}^{J} n_j \hat{\mu}_j$$

- We define the between-class variance matrix as:

$$\mathbf{S}_B = \sum_{j=1}^{J} n_j \left(\hat{\mu}_j - \hat{\mu}\right) \left(\hat{\mu}_j - \hat{\mu}\right)^\top$$

and the within-class variance matrix as:

$$\mathbf{S}_W = \sum_{j=1}^{J} \mathbf{S}_j$$

- We can show that the total variance matrix can be decomposed into the sum of the within-class and between-class variance matrices:

$$\mathbf{S} = \mathbf{S}_W + \mathbf{S}_B$$

Discriminant analysis

Class separation maximization

- The discriminant analysis consists in finding the discriminant linear combination $\beta^\top X$ that has the maximum between-class variance relative to the within-class variance:

$$\beta^\star = \arg\max \mathcal{J}(\beta)$$

where $\mathcal{J}(\beta)$ is the Fisher criterion:

$$\mathcal{J}(\beta) = \frac{\beta^\top \mathbf{S}_B \beta}{\beta^\top \mathbf{S}_W \beta}$$

- Since the objective function is invariant if we rescale the vector β – $\mathcal{J}(\beta') = \mathcal{J}(\beta)$ if $\beta' = c\beta$, we can impose that $\beta^\top \mathbf{S}_W \beta = 1$. It follows that:

$$\begin{aligned} \hat{\beta} &= \arg\max \beta^\top \mathbf{S}_B \beta \\ \text{s.t.} &\quad \beta^\top \mathbf{S}_W \beta = 1 \end{aligned}$$

Discriminant analysis

Class separation maximization

- The Lagrange function is:

$$\mathcal{L}(\beta;\lambda) = \beta^\top \mathbf{S}_B \beta - \lambda \left(\beta^\top \mathbf{S}_W \beta - 1\right)$$

- We deduce that the first-order condition is equal to:

$$\frac{\partial\, \mathcal{L}(\beta;\lambda)}{\partial\, \beta^\top} = 2\mathbf{S}_B\beta - 2\lambda\mathbf{S}_W\beta = \mathbf{0}$$

- It is remarkable that we obtain a generalized eigenvalue $\mathbf{S}_B\beta = \lambda\mathbf{S}_W\beta$ or equivalently:

$$\mathbf{S}_W^{-1}\mathbf{S}_B\beta = \lambda\beta$$

- Even if $\mathbf{S}_W$ and $\mathbf{S}_B$ are two symmetric matrices, it is not necessarily the case for the product $\mathbf{S}_W^{-1}\mathbf{S}_B$
- Using the eigendecomposition $\mathbf{S}_B = V\Lambda V^\top$, we have $\mathbf{S}_B^{1/2} = V\Lambda^{1/2}V^\top$

Discriminant analysis

Class separation maximization

- With the parametrization $\alpha = \mathbf{S}_B^{1/2}\beta$, the first-order condition becomes:
$$\mathbf{S}_B^{1/2}\mathbf{S}_W^{-1}\mathbf{S}_B^{1/2}\alpha = \lambda\alpha$$
because $\beta = \mathbf{S}_B^{-1/2}\alpha$
- We have a right regular eigenvalue problem
- Let λ_k and v_k be the k^{th} eigenvalue and eigenvector of the symmetric matrix $\mathbf{S}_B^{1/2}\mathbf{S}_W^{-1}\mathbf{S}_B^{1/2}$
- It is obvious that the optimal solution $\alpha^\star$ is the first eigenvector v_1 corresponding to the largest eigenvalue λ_1
- We conclude that the estimator is $\hat{\beta} = \mathbf{S}_B^{-1/2}v_1$ and the discriminant linear relationship is $Y^c = v_1^\top \mathbf{S}_B^{-1/2}X$
- Moreover, we have:
$$\lambda_1 = J\left(\hat{\beta}\right) = \frac{\hat{\beta}^\top \mathbf{S}_B\hat{\beta}}{\hat{\beta}^\top \mathbf{S}_W\hat{\beta}}$$

Discriminant analysis
Class separation maximization

Example #3

We consider a problem with two classes $\mathcal{C}_1$ and $\mathcal{C}_2$, and two explanatory variables (X_1, X_2). Class $\mathcal{C}_1$ is composed of 7 observations: $(1,2)$, $(1,4)$, $(3,6)$, $(3,3)$, $(4,2)$, $(5,6)$, $(5,5)$, whereas class $\mathcal{C}_2$ is composed of 6 observations: $(1,0)$, $(2,1)$, $(4,1)$, $(3,2)$, $(6,4)$ and $(6,5)$.

Discriminant analysis
Class separation maximization

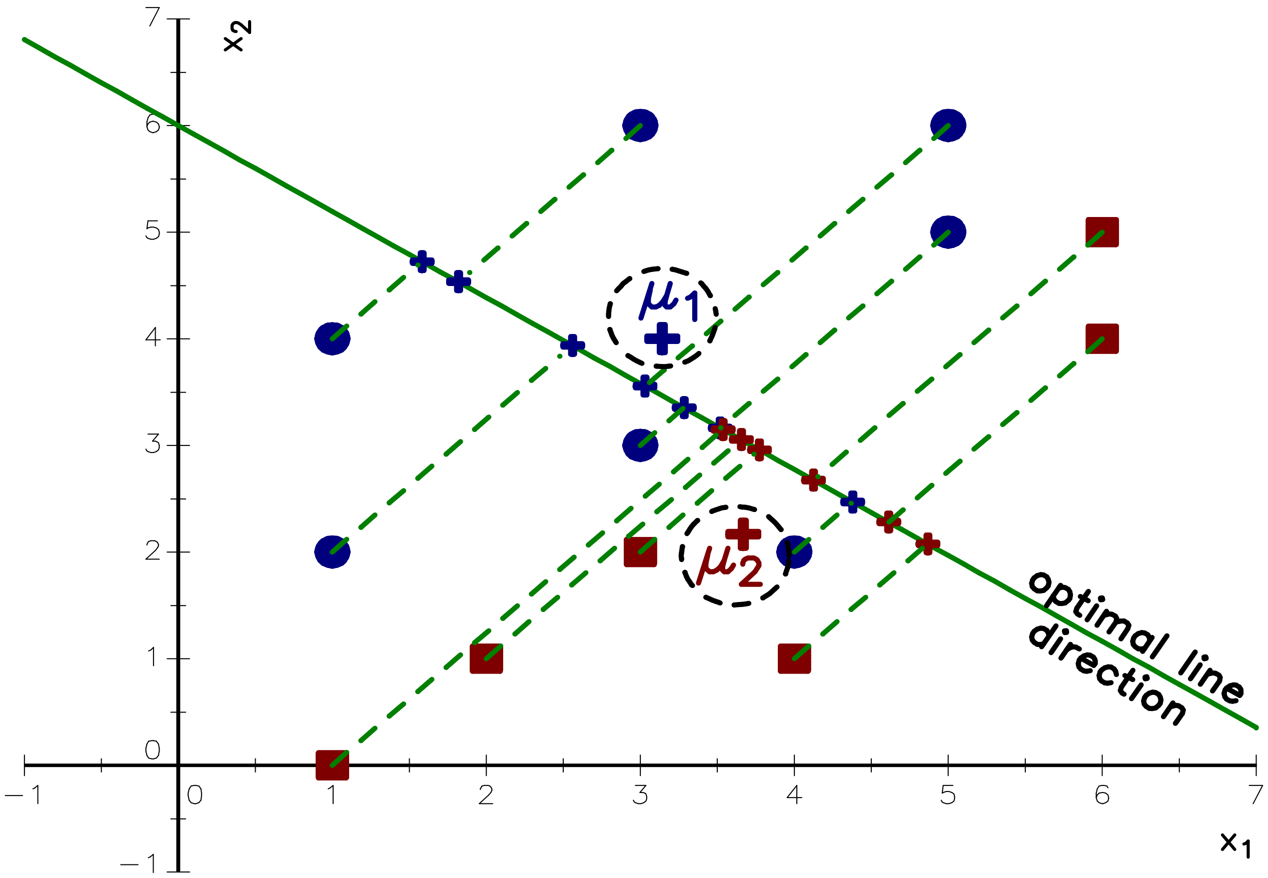


Figure: Linear projection and the Fisher solution

Discriminant analysis

Class separation maximization

Concerning the assignment decision, we can consider the midpoint rule:

$$\left\{ \begin{array}{l} s_i < \bar{\mu} \Rightarrow i \in \mathcal{C}_1 \\ s_i > \bar{\mu} \Rightarrow i \in \mathcal{C}_2 \end{array} \right.$$

where $\bar{\mu} = \left(\bar{\mu}_1 + \bar{\mu}_2\right)/2$, $\bar{\mu}_1 = \beta^\top \hat{\mu}_1$ and $\bar{\mu}_2 = \beta^\top \hat{\mu}_2$

This rule is not optimal because it does not depend on the variance $\bar{s}_1^2$ and $\bar{s}_2^2$ of each class

Discriminant analysis

Class separation maximization

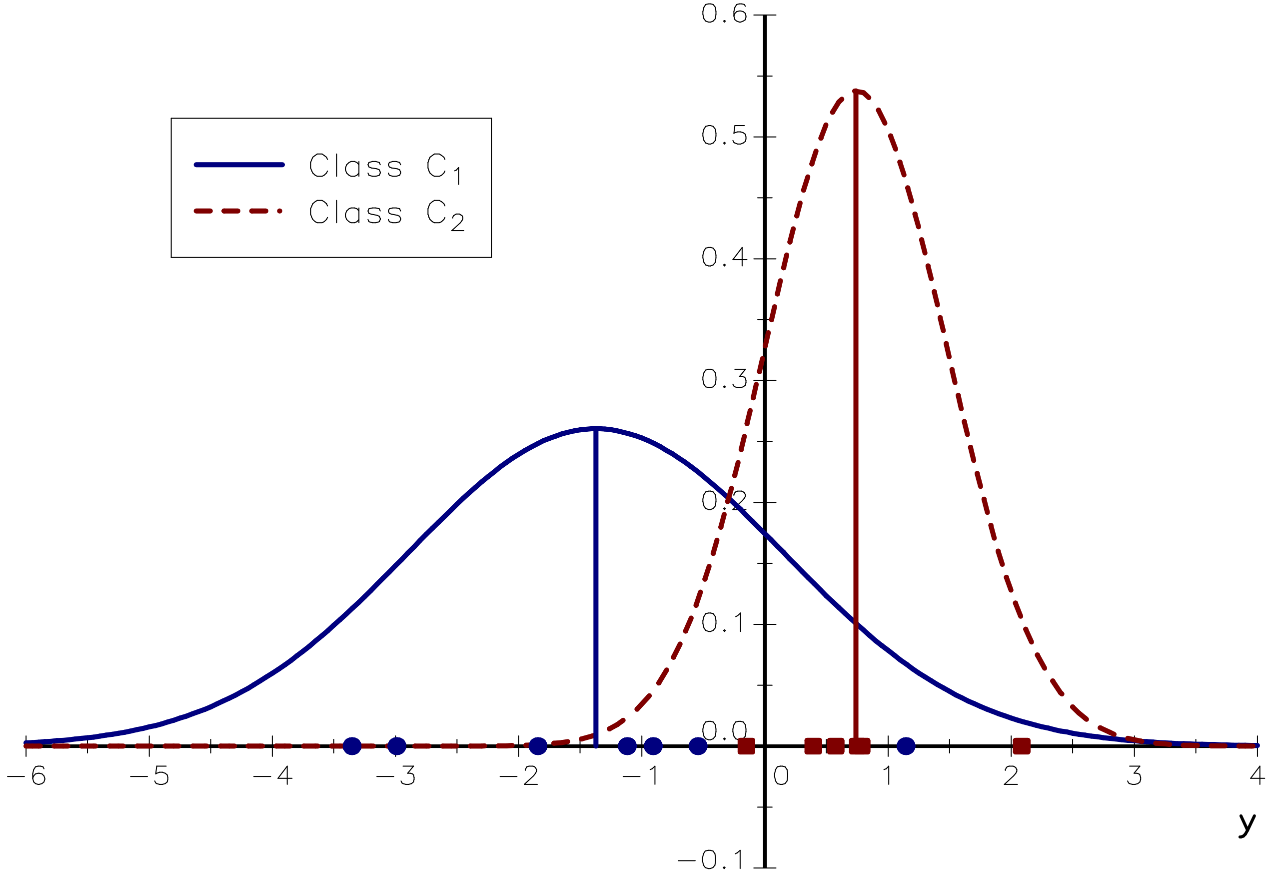


Figure: Class separation and the cut-off criterion

Binary choice models

General framework

- We assume that Y can take two values 0 and 1
- We consider models that link the outcome to a set of factors X:

$$\Pr\{Y = 1 \mid X = x\} = \mathbf{F}\left(x^\top \beta\right)$$

- $\mathbf{F}$ must be a cumulative distribution function in order to ensure that $\mathbf{F}(z) \in [0, 1]$
- We also assume that the model is symmetric, implying that $\mathbf{F}(z) + \mathbf{F}(-z) = 1$
- Given a sample $\{(x_i, y_i), i = 1, \ldots, n\}$, the log-likelihood function is equal to:

$$\ell(\theta) = \sum_{i=1}^{n} \ln \Pr\{Y_i = y_i\}$$

where y_i takes the values 0 or 1

Binary choice models

General framework

- We have:

$$\Pr\left\{ Y_i = y_i \right\} = p_i^{y_i} \cdot (1 - p_i)^{1 - y_i}$$

  where $p_i = \Pr\left\{ Y_i = 1 \mid X_i = x_i \right\}$

- We deduce that:

$$\begin{aligned} \ell(\theta) &= \sum_{i=1}^{n} y_i \ln p_i + (1 - y_i) \ln (1 - p_i) \\ &= \sum_{i=1}^{n} y_i \ln \mathbf{F}\left(x_i^\top \beta\right) + (1 - y_i) \ln \left(1 - \mathbf{F}\left(x_i^\top \beta\right)\right) \end{aligned}$$

- We notice that the vector θ includes only the parameters β

Binary choice models

General framework

- By noting $f(z)$ the probability density function, it follows that the associated score vector of the log-likelihood function is:

$$\begin{aligned}\mathcal{S}(\beta) &= \frac{\partial\,\ell(\beta)}{\partial\,\beta} \\ &= \sum_{i=1}^{n} \frac{f\left(x_i^\top \beta\right)}{\mathbf{F}\left(x_i^\top \beta\right)\mathbf{F}\left(-x_i^\top \beta\right)} \left(y_i - \mathbf{F}\left(x_i^\top \beta\right)\right) x_i\end{aligned}$$

Binary choice models

General framework

- The Hessian matrix is:

$$H(\beta) = \frac{\partial^2 \boldsymbol{\ell}(\beta)}{\partial \beta \, \partial \beta^\top} = -\sum_{i=1}^{n} H_i \cdot \left(x_i x_i^\top\right)$$

where:

$$\begin{aligned} H_i &= \frac{f\left(x_i^\top \beta\right)^2}{\mathbf{F}\left(x_i^\top \beta\right) \mathbf{F}\left(-x_i^\top \beta\right)} - \left(y_i - \mathbf{F}\left(x_i^\top \beta\right)\right) \cdot \\ & \left(\frac{f'\left(x_i^\top \beta\right)}{\mathbf{F}\left(x_i^\top \beta\right) \mathbf{F}\left(-x_i^\top \beta\right)} - \frac{f\left(x_i^\top \beta\right)^2 \left(1 - 2\mathbf{F}\left(x_i^\top \beta\right)\right)}{\mathbf{F}\left(x_i^\top \beta\right)^2 \mathbf{F}\left(-x_i^\top \beta\right)^2}\right) \end{aligned}$$

Binary choice models

General framework

- Once $\hat{\beta}$ is estimated by the method of maximum likelihood, we can calculated the predicted probability for the i^{th} observation:

$$\hat{p}_i = \mathbf{F}\left(x_i^\top \hat{\beta}\right)$$

- Like a linear regression model, we can define the residual as the difference between the observation y_i and the predicted value $\hat{p}_i$
- We can also exploit the property that the conditional distribution of Y_i is a Bernoulli distribution $\mathcal{B}(p_i)$
- It is better to use the standardized (or Pearson) residuals:

$$\hat{u}_i = \frac{y_i - \hat{p}_i}{\sqrt{\hat{p}_i (1 - \hat{p}_i)}}$$

- These residuals are related to the Pearson's chi-squared statistic:

$$\chi^2_{\mathrm{Pearson}} = \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{p}_i)^2}{\hat{p}_i (1 - \hat{p}_i)}$$

Binary choice models

General framework

- This statistic may used to measure the goodness-of-fit of the model
- Under the assumption $\mathcal{H}_0$ that there is no lack-of-fit, we have $\chi^2_{\text{Pearson}} \sim \chi^2_{n-K}$ where K is the number of exogenous variables
- Another goodness-of-fit statistic is the likelihood ratio. For the '*saturated*' model, the estimated probability $\hat{p}_i$ is exactly equal to y_i
- We deduce that the likelihood ratio is equal to:

$$-2\ln\Lambda = 2\sum_{i=1}^{n} y_i \ln\left(\frac{y_i}{\hat{p}_i}\right) + (1-y_i)\ln\left(\frac{1-y_i}{1-\hat{p}_i}\right)$$

Binary choice models

General framework

- In binomial choice models, $D = -2 \ln \Lambda$ is also called the deviance and we have $D \sim \chi^2_{n-K}$
- In a perfect fit $\hat{p}_i = y_i$, the likelihood ratio is exactly equal to zero
- The forecasting procedure consists of estimating the probability $\hat{p} = \mathbf{F}\left(x^\top \hat{\beta}\right)$ for a given set of variables x and to use the following decision criterion:

$$Y = 1 \Leftrightarrow \hat{p} \geq \frac{1}{2}$$

Binary choice models

Logistic regression

- The logit model uses the following cumulative distribution function:

$$\mathbf{F}(z) = \frac{1}{1+e^{-z}} = \frac{e^{z}}{e^{z}+1}$$

- The probability density function is then equal to:

$$f(z) = \frac{e^{-z}}{\left(1+e^{-z}\right)^2}$$

- The log-likelihood function is equal to:

$$\begin{aligned}
\ell(\beta) &= \sum_{i=1}^{n} (1-y_i)\ln\left(1-\mathbf{F}\left(x_i^\top\beta\right)\right) + y_i \ln \mathbf{F}\left(x_i^\top\beta\right) \\
&= \sum_{i=1}^{n} (1-y_i)\ln\left(\frac{e^{-x_i^\top\beta}}{1+e^{-x_i^\top\beta}}\right) - y_i \ln\left(1+e^{-x_i^\top\beta}\right) \\
&= -\sum_{i=1}^{n} \ln\left(1+e^{-x_i^\top\beta}\right) + (1-y_i)\left(x_i^\top\beta\right)
\end{aligned}$$

Binary choice models

Logistic regression

- We also have:

$$\mathcal{S}(\beta) = \sum_{i=1}^{n} \left(y_i - \mathbf{F}\left(x_i^\top \beta \right) \right) x_i$$

and:

$$H(\beta) = -\sum_{i=1}^{n} f\left(x_i^\top \beta \right) \cdot \left(x_i x_i^\top \right)$$

Binary choice models

Probit analysis

- The probit model assumes that $\mathbf{F}(z)$ is the Gaussian distribution
- The log-likelihood function is then:

$$\ell(\beta) = \sum_{i=1}^{n} (1 - y_i) \ln \left(1 - \Phi\left(x_i^\top \beta\right)\right) + y_i \ln \Phi\left(x_i^\top \beta\right)$$

- The probit model can be seen as a latent variable model
- Let us consider the linear model $Y^\star = \beta^\top X + U$ where $U \sim \mathcal{N}\left(0, \sigma^2\right)$
- We assume that we do not observe $Y^\star$ but $Y = g\left(Y^\star\right)$
- For example, if $g(z) = \mathbb{1}\{z > 0\}$, we obtain:

$$\Pr\{Y = 1 \mid X = x\} = \Pr\left\{\beta^\top X + U > 0 \mid X = x\right\} = \Phi\left(\frac{\beta^\top x}{\sigma}\right)$$

- We notice that only the ratio β/σ is identifiable
- Since we can set $\sigma = 1$, we obtain the probit model

Binary choice models

Regularization

- The regularized log-likelihood function is equal to:

$$\ell(\theta; \lambda) = \ell(\theta) - \frac{\lambda}{p} \|\theta\|_p^p$$

- The case $p = 1$ is equivalent to consider a lasso penalization
- The case $p = 2$ corresponds to the ridge regularization

Binary choice models

Extension to multinomial logistic regression

- We assume that Y can take J labels $(\mathfrak{L}_1, \ldots, \mathfrak{L}_J)$ or belongs to J disjoint classes $(\mathcal{C}_1, \ldots, \mathcal{C}_J)$
- We define the conditional probability as follows:

$$p_j(x) = \Pr\{Y = \mathfrak{L}_j \mid X = x\} = \Pr\{Y \in \mathcal{C}_j \mid X = x\} = \frac{e^{\beta_j^\top x}}{1 + \sum_{j=1}^{J-1} e^{\beta_j^\top x}}$$

- The probability of the last label is then equal to:

$$p_J(x) = 1 - \sum_{j=1}^{J-1} p_j(x) = \frac{1}{1 + \sum_{j=1}^{J-1} e^{\beta_j^\top x}}$$

- The log-likelihood function becomes:

$$\ell(\theta) = \sum_{i=1}^{n} \ln \left(\prod_{j=1}^{J} p_j(x_i)^{i \in \mathcal{C}_j} \right)$$

where θ is the vector of parameters $(\beta_1, \ldots, \beta_{J-1})$

Non-parametric supervised methods

- k-nearest neighbor classifier (k-NN)
- Neural networks (NN)
- Support vector machines (SVM)
- Model averaging (bagging or bootstrap aggregation, random forests, boosting)

Definition and properties

- The entropy is a measure of unpredictability or uncertainty of a random variable
- Let (X, Y) be a random vector where $p_{i,j} = \Pr\{X = x_i, Y = y_j\}$, $p_i = \Pr\{X = x_i\}$ and $p_j = \Pr\{Y = y_j\}$
- The Shannon entropy of the discrete random variable X is given by:

$$H(X) = -\sum_{i=1}^{n} p_i \ln p_i$$

- We have the property $0 \leq H(X) \leq \ln n$.
- The Shannon entropy is a measure of the average information of the system
- The lower the Shannon entropy, the more informative the system

Definition and properties

- For a random vector (X, Y), we have:

$$H(X, Y) = -\sum_{i=1}^{n} \sum_{j=1}^{n} p_{i,j} \ln p_{i,j}$$

- We deduce that the conditional information of Y given X is equal to:

$$\begin{aligned} H(Y \mid X) &= \mathbb{E}_X \left[H(Y \mid X = x) \right] \\ &= -\sum_{i=1}^{n} \sum_{j=1}^{n} p_{i,j} \ln \frac{p_{i,j}}{p_i} \\ &= H(X, Y) - H(X) \end{aligned}$$

Definition and properties

We have the following properties:

- if X and Y are independent, we have $H(Y \mid X) = H(Y)$ and $H(X,Y) = H(Y) + H(X)$;
- if X and Y are perfectly dependent, we have $H(Y \mid X) = 0$ and $H(X,Y) = H(X)$.

The amount of information obtained about one random variable, through the other random variable is measured by the mutual information:

$$
\begin{aligned}
I(X,Y) &= H(Y) + H(X) - H(X,Y) \\
&= \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} p_{i,j} \ln \frac{p_{i,j}}{p_i p_j}
\end{aligned}
$$

Definition and properties

$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$

$$H(X) = H(Y) = 1.792$$
$$H(X, Y) = 3.584$$
$$I(X, Y) = 0$$

$1/6$					
	$1/6$				
		$1/6$			
			$1/6$		
				$1/6$	
					$1/6$

$$H(X) = H(Y) = 1.792$$
$$H(X, Y) = 1.792$$
$$I(X, Y) = 1.792$$

Figure: Examples of Shannon entropy calculation

Definition and properties

1/24	1/24				
1/24	1/24	1/24	1/48		
	1/24	1/6	1/24	1/48	
	1/48	1/24	1/6	1/24	
		1/48	1/24	1/24	1/24
				1/24	1/24

$$H(X) = H(Y) = 1.683$$
$$H(X,Y) = 2.774$$
$$I(X,Y) = 0.593$$

					1/12
1/8			1/8		
	1/24				
5/24		1/24			
3/24				1/24	
3/24	1/24	1/24			

$$H(X) = 1.658$$
$$H(Y) = 1.328$$
$$I(X,Y) = 0.750$$

Figure: Examples of Shannon entropy calculation

Application to scoring

- Let S and Y be the score and the control variable
- For instance, Y is a binary random variable that may indicate a bad credit ($Y = 0$) or a good credit ($Y = 1$)
- We consider the following decision rule:

$$\begin{cases} S \leq 0 \Rightarrow S^{\star} = 0 \\ S > 0 \Rightarrow S^{\star} = 1 \end{cases}$$

Application to scoring

- We note $n_{i,j}$ the number of observations such that $S^\star = i$ and $Y = j$. We obtain the following system $(S^\star, Y)$:

	$Y = 0$	$Y = 1$
$S^\star = 0$	$n_{0,0}$	$n_{0,1}$
$S^\star = 1$	$n_{1,0}$	$n_{1,1}$

  where $n = n_{0,0} + n_{0,1} + n_{1,0} + n_{1,1}$ is the total number of observations
- The hit rate is the ratio of good bets:

$$H = \frac{n_{0,0} + n_{1,1}}{n}$$

- This statistic can be viewed as an information measure of the system (S, Y)
- When there are more states, we can consider the Shannon entropy

Application to scoring

	y_1	y_2	y_3	y_4	y_5
s_1	10	9			
s_2	7	9			
s_3	3		7	2	
s_4		2	10	4	5
s_5				10	2
s_6			3	4	13

$$H(S_1) = 1.767$$
$$H(Y) = 1.609$$
$$H(S_1, Y) = 2.614$$
$$I(S_1, Y) = 0.763$$

	y_1	y_2	y_3	y_4	y_5
s_1	7	10			
s_2	10	8			
s_3			5	4	3
s_4	3		10	6	4
s_5	2			5	8
s_6			5	5	5

$$H(S_1) = 1.771$$
$$H(Y) = 1.609$$
$$H(S_1, Y) = 2.745$$
$$I(S_1, Y) = 0.636$$

Figure: Scorecards S_1 and S_2

Graphical methods

- We assume that the control variable Y can takes two values
  - $Y = 0$ corresponds to a bad risk (or bad signal)
  - $Y = 1$ corresponds to a good risk (or good signal)

Graphical methods

- We assume that the probability $\Pr\{Y = 1 \mid S \geq s\}$ is increasing with respect to the level $s \in [0, 1]$, which corresponds to the rate of acceptance.
- We deduce that the decision rule is the following:
  - if the score of the observation is above the threshold s, the observation is selected;
  - if the score of the observation is below the threshold s, the observation is not selected.
- If s is equal to one, we select no observation
- If s is equal to zero, we select all the observations

Performance curve

- The performance curve is the parametric function $y = \mathcal{P}(x)$ defined by:

$$\begin{cases} x(s) = \Pr\{S \geq s\} \\ y(s) = \dfrac{\Pr\{Y = 0 \mid S \geq s\}}{\Pr\{Y = 0\}} \end{cases}$$

  where $x(s)$ corresponds to the proportion of selected observations and $y(s)$ corresponds to the ratio between the proportion of selected bad risks and the proportion of bad risks in the population
- The score is efficient if the ratio is below one
- If $y(s) > 1$, the score selects more bad risks than those we can find in the population
- If $y(s) = 1$, the score is random and the performance is equal to zero. In this case, the selected population is representative of the total population

Selection curve

- The selection curve is the parametric curve $y = \mathcal{S}(x)$ defined by:

$$\begin{cases} x(s) = \Pr\{S \geq s\} \\ y(s) = \Pr\{S \geq s \mid Y = 0\} \end{cases}$$

  where $y(s)$ corresponds to the ratio of observations that are wrongly selected

- By construction, we would like that the curve $y = \mathcal{S}(x)$ is located below the bisecting line $y = x$ in order to verify that $\Pr\{S \geq s \mid Y = 0\} < \Pr\{S \geq s\}$

Performance and selection curves

- We have:

$$\begin{aligned}\Pr\{S \geq s \mid Y = 0\} &= \frac{\Pr\{S \geq s, Y = 0\}}{\Pr\{Y = 0\}} \\ &= \Pr\{S \geq s\} \cdot \frac{\Pr\{S \geq s, Y = 0\}}{\Pr\{S \geq s\}\Pr\{Y = 0\}} \\ &= \Pr\{S \geq s\} \cdot \frac{\Pr\{Y = 0 \mid S \geq s\}}{\Pr\{Y = 0\}}\end{aligned}$$

- The performance and selection curves are related as follows:

$$\mathcal{S}(x) = x\mathcal{P}(x)$$

Discriminant curve

- The discriminant curve is the parametric curve $y = \mathcal{D}(x)$ defined by:

$$\mathcal{D}(x) = g_1\left(g_0^{-1}(x)\right)$$

where:

$$g_y(s) = \Pr\{S \geq s \mid Y = y\}$$

- It represents the proportion of good risks in the selected population with respect to the proportion of bad risks in the selected population
- The score is said to be discriminant if the curve $y = \mathcal{D}(x)$ is located above the bisecting line $y = x$

Some properties

1. the performance curve (respectively, the selection curve) is located below the line $y = 1$ (respectively, the bisecting line $y = x$) if and only if $\operatorname{cov}\left(f\left(Y\right), g\left(S\right)\right) \geq 0$ for any increasing functions f and g
2. the performance curve is increasing if and only if:

   $$\operatorname{cov}\left(f\left(Y\right), g\left(S\right) \mid S \geq s\right) \geq 0$$

   for any increasing functions f and g, and any threshold level s
3. the selection curve is convex if and only if $\mathbb{E}\left[f\left(Y\right) \mid S = s\right]$ is increasing with respect to the threshold level s for any increasing function f
4. We can show that (3) $\Rightarrow$ (2) $\Rightarrow$ (1)

Some properties

- A score is perfect or optimal if there is a threshold level $s^\star$ such that $\Pr\{Y = 1 \mid S \geq s^\star\} = 1$ and $\Pr\{Y = 0 \mid S < s^\star\} = 1$
- It separates the population between good and bad risks
- Graphically, the selection curve of a perfect score is equal to:

$$y = \mathbb{1}\{x > \Pr\{Y = 1\}\} \cdot \left(1 + \frac{x - 1}{\Pr\{Y = 0\}}\right)$$

- Using the relationship $\mathcal{S}(x) = x\mathcal{P}(x)$, we deduce that the performance curve of a perfect score is given by:

$$y = \mathbb{1}\{x > \Pr\{Y = 1\}\} \cdot \left(\frac{x - \Pr\{Y = 1\}}{x \cdot \Pr\{Y = 0\}}\right)$$

- For the discriminant curve, a perfect score satisfies $\mathcal{D}(x) = 1$
- When the score is random, we have $\mathcal{S}(x) = \mathcal{D}(x) = x$ and $\mathcal{P}(x) = 1$

Some properties

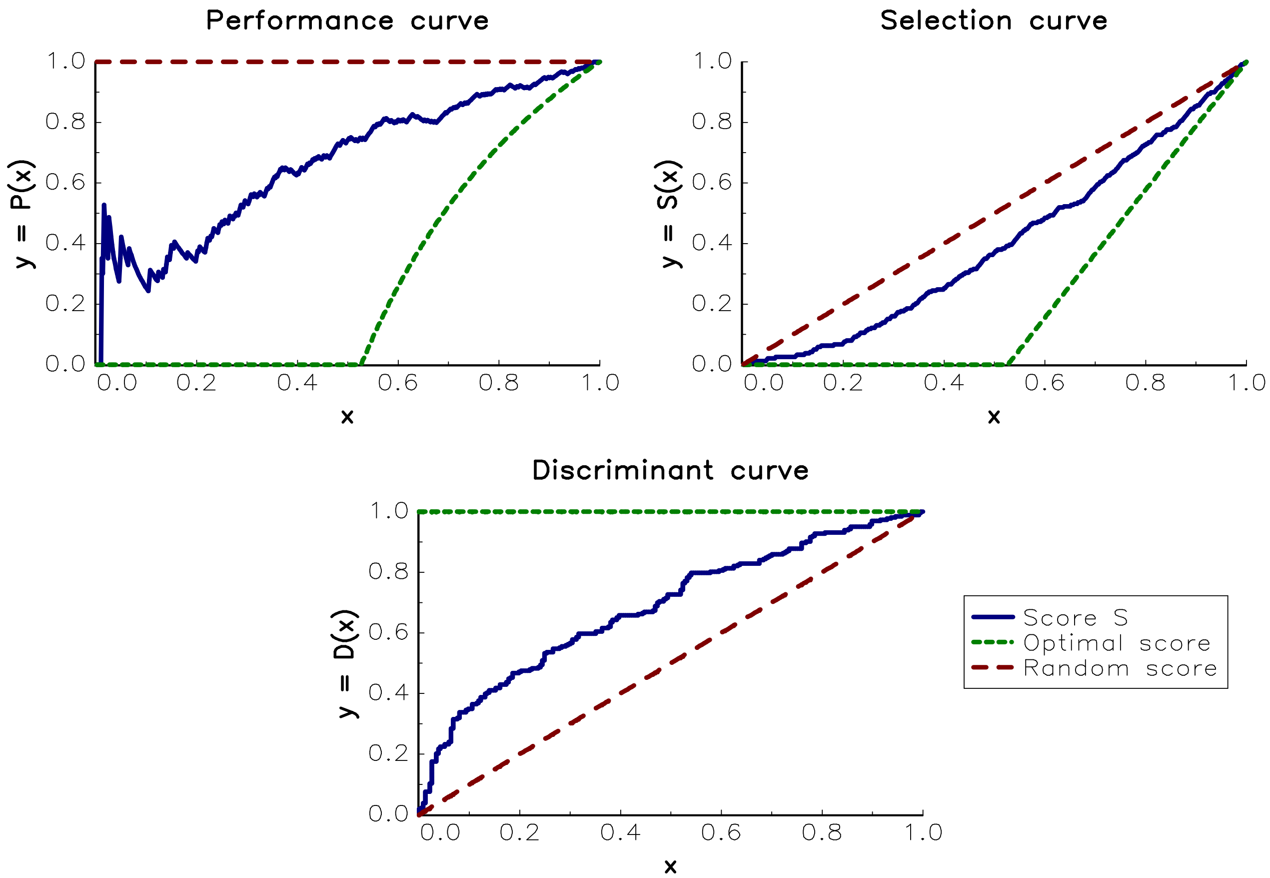


Figure: Performance, selection and discriminant curves

Some properties

- The score S_1 is more performing on the population P_1 than the score S_2 on the population P_2 if and only if the performance (or selection) curve of (S_1, P_1) is below the performance (or selection) curve of (S_2, P_2)
- The score S_1 is more discriminatory on the population P_1 than the score S_2 on the population P_2 if and only if the discriminant curve of (S_1, P_1) is above the discriminant curve of (S_2, P_2)

Some properties

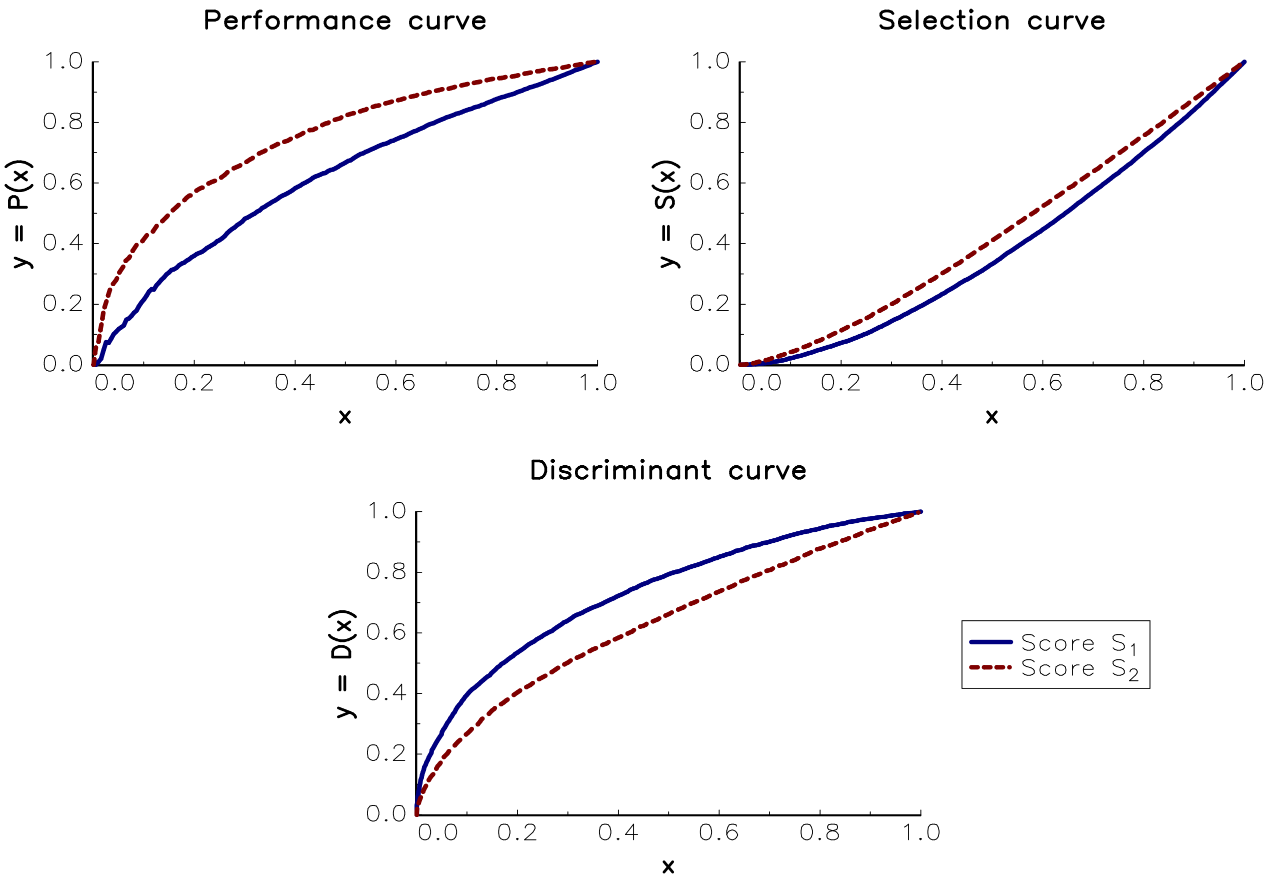


Figure: The score S_1 is better than the score S_2

Some properties

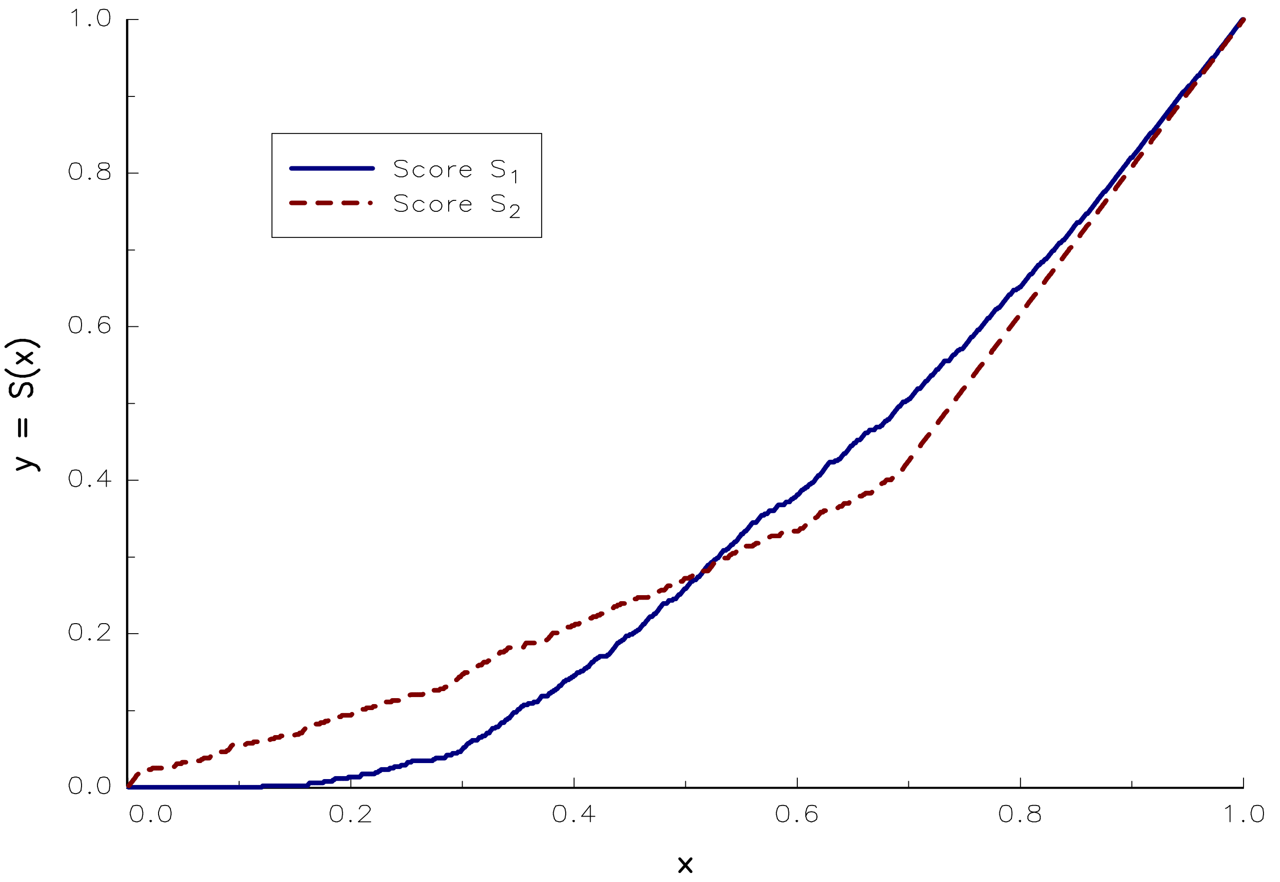


Figure: Illustration of the partial ordering between two scores

Kolmogorov-Smirnov test

- We consider the cumulative distribution functions:

$$\mathbf{F}_0(s) = \Pr\{S \leq s \mid Y = 0\}$$

  and:

$$\mathbf{F}_1(s) = \Pr\{S \leq s \mid Y = 1\}$$

- The score S is relevant if we have the stochastic dominance order $\mathbf{F}_0 \succ \mathbf{F}_1$
- In this case, the score quality is measured by the Kolmogorov-Smirnov statistic:

$$\mathrm{KS} = \max_s |\mathbf{F}_0(s) - \mathbf{F}_1(s)|$$

  It takes the value 1 if the score is perfect
- The KS statistic may be used to verify that the score is not random ($\mathcal{H}_0 : \mathrm{KS} = 0$)

Kolmogorov-Smirnov test

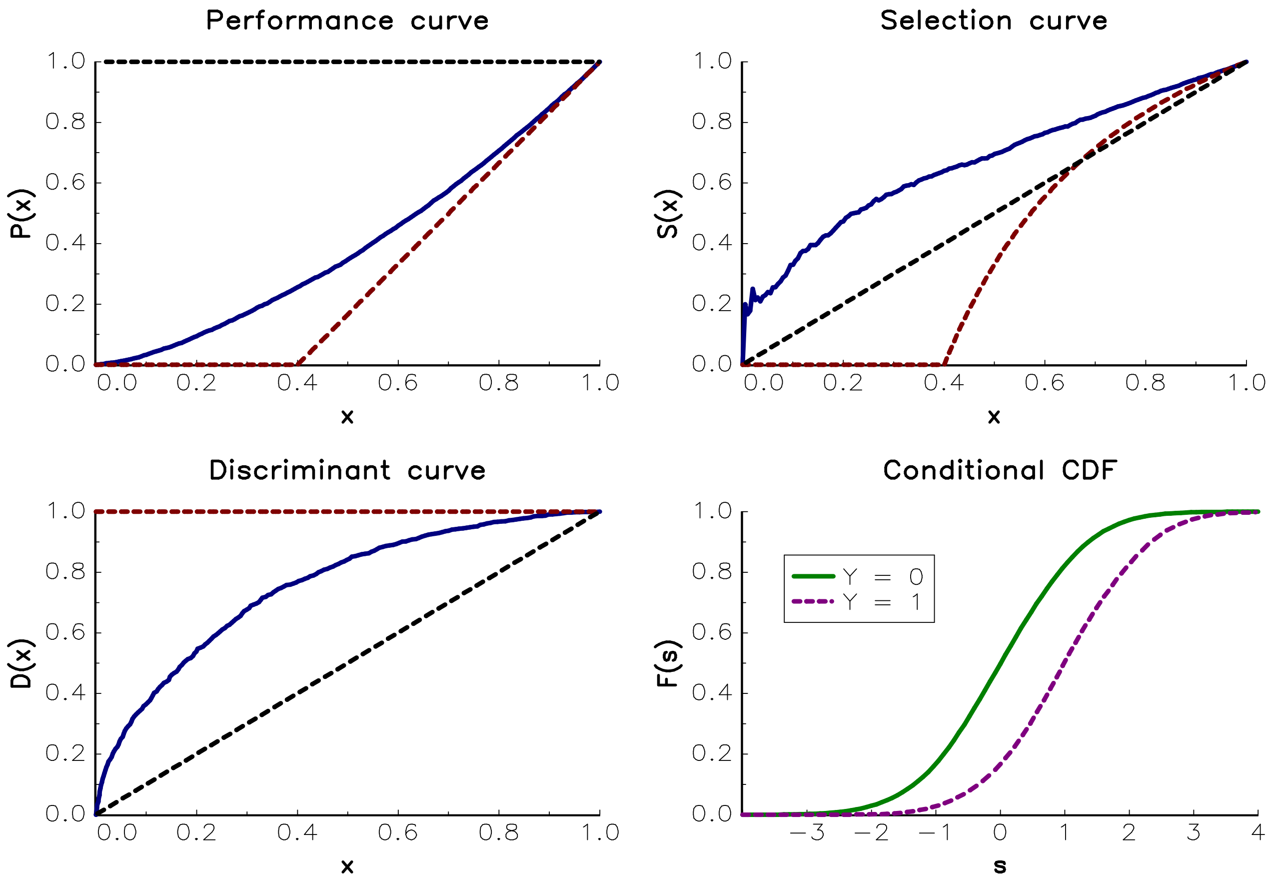


Figure: Comparison of the distributions $\mathbf{F}_0(s)$ and $\mathbf{F}_1(s)$

Gini coefficient

The Lorenz curve

- Let X and Y be two random variables
- The Lorenz curve $y = \mathcal{L}(x)$ is the parametric curve defined by:

$$\left\{ \begin{array}{l} x = \Pr\{X \leq x\} \\ y = \Pr\{Y \leq y \mid X \leq x\} \end{array} \right.$$

- In economics, x represents the proportion of individuals that are ranked by income while y represents the proportion of income
- In this case, the Lorenz curve is a graphical representation of the distribution of income and is used for illustrating inequality of the wealth distribution between individuals

Gini coefficient

The Lorenz curve

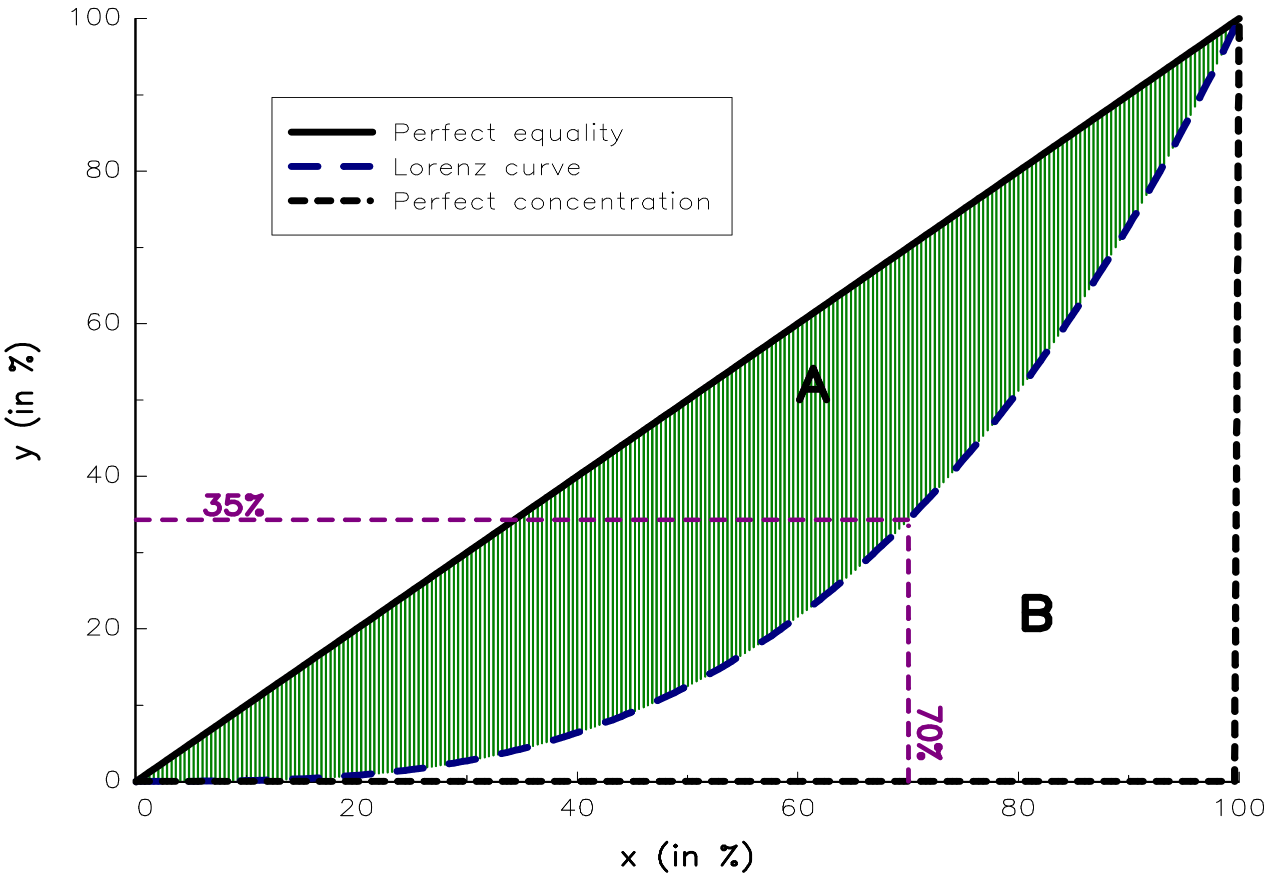


Figure: An example of Lorenz curve

Gini coefficient
Definition

- We define the Gini coefficient by:

$$\mathcal{G}ini\left(\mathcal{L}\right) = \frac{A}{A+B}$$

  where A is the area between the Lorenz curve and the curve of perfect equality, and B is the area between the curve of perfect concentration and the Lorenz curve
- By construction, we have $0 \leq \mathcal{G}ini\left(\mathcal{L}\right) \leq 1$
- The Gini coefficient is equal to zero in the case of perfect equality and one in the case of perfect concentration
- We have:

$$\mathcal{G}ini\left(\mathcal{L}\right) = 1 - 2\int_0^1 \mathcal{L}\left(x\right)\,\mathrm{d}x$$

Gini coefficient

Application to credit scoring

- We can interpret the selection curve as a Lorenz curve
- We recall that $\mathbf{F}(s) = \Pr\{S \leq s\}$, $\mathbf{F}_0(s) = \Pr\{S \leq s \mid Y = 0\}$ and $\mathbf{F}_1(s) = \Pr\{S \leq s \mid Y = 1\}$
- The selection curve is defined by the following parametric coordinates:

$$\left\{ \begin{array}{l} x(s) = 1 - \mathbf{F}(s) \\ y(s) = 1 - \mathbf{F}_0(s) \end{array} \right.$$

- The selection curve measures the capacity of the score for not selecting bad risks
- We could also build the Lorenz curve that measures the capacity of the score for selecting good risks:

$$\left\{ \begin{array}{l} x(s) = \Pr\{S \geq s\} = 1 - \mathbf{F}(s) \\ y(s) = \Pr\{S \geq s \mid Y = 1\} = 1 - \mathbf{F}_1(s) \end{array} \right.$$

It is called the precision curve

Gini coefficient

Application to credit scoring

- Another popular graphical tool is the receiver operating characteristic (or ROC curve), which is defined by:

$$\left\{ \begin{array}{l} x(s) = \Pr\{S \geq s \mid Y = 0\} = 1 - \mathbf{F}_0(s) \\ y(s) = \Pr\{S \geq s \mid Y = 1\} = 1 - \mathbf{F}_1(s) \end{array} \right.$$

- The Gini coefficient associated to the Lorenz curve $\mathcal{L}$ becomes:

$$\mathcal{G}ini(\mathcal{L}) = 2\int_0^1 \mathcal{L}(x)\,\mathrm{d}x - 1$$

- The Gini coefficient of the score S is then computed as follows:

$$\mathcal{G}ini^\star(S) = \frac{\mathcal{G}ini(\mathcal{L})}{\mathcal{G}ini(\mathcal{L}^\star)}$$

where $\mathcal{L}^\star$ is the Lorenz curve associated to the perfect score

- An alternative to the Gini coefficient is the AUC measure, which corresponds to the area under the ROC curve:

$$\mathcal{G}ini(\mathrm{ROC}) = 2 \times \mathrm{AUC}(\mathrm{ROC}) - 1$$

Gini coefficient

Application to credit scoring

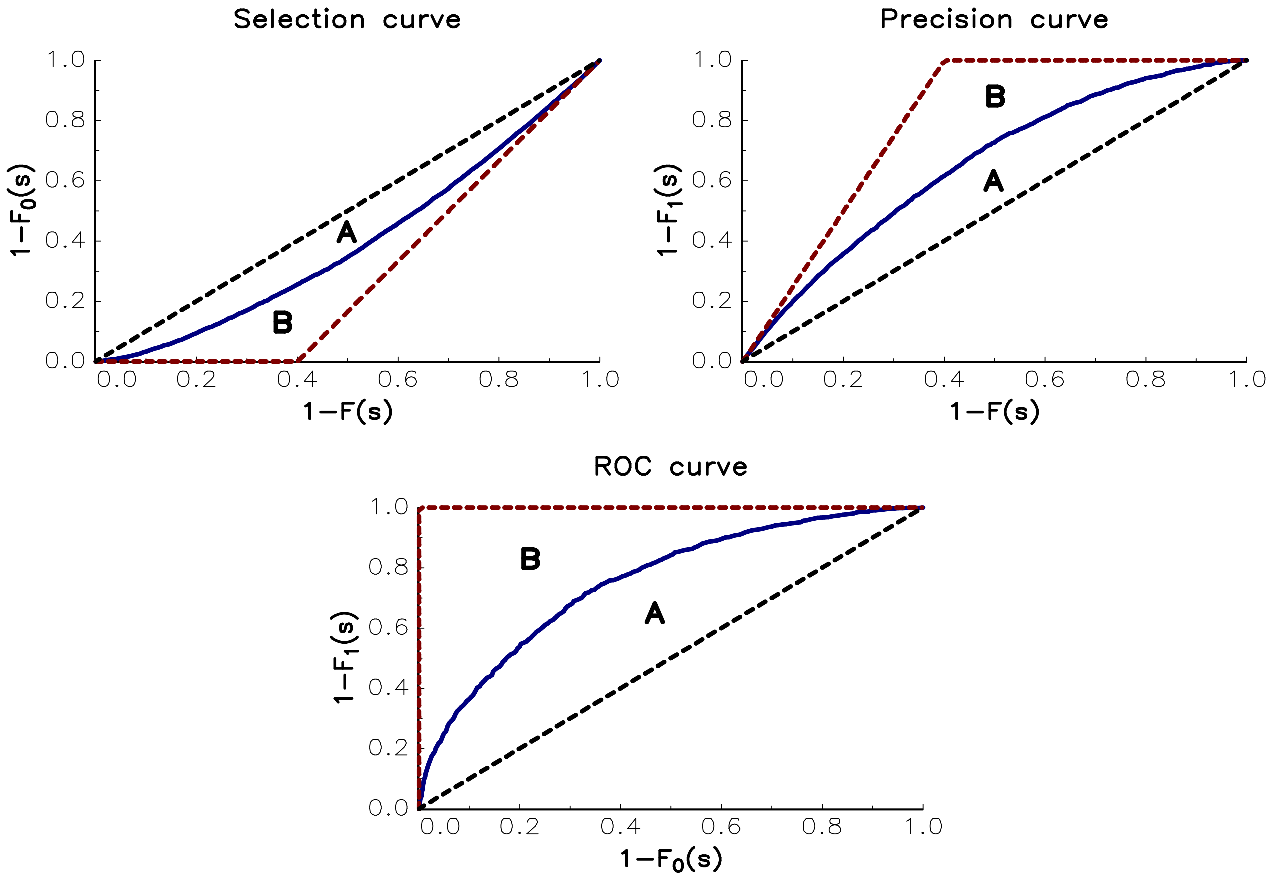


Figure: Selection, precision and ROC curves

Choice of the optimal cut-off

Confusion matrix

- A confusion matrix is a special case of contingency matrix
- Each row of the matrix represents the frequency in a predicted class while each column represents the frequency in an actual class
- Using the test set, it takes the following form:

	$Y = 0$	$Y = 1$
$S < s$	$n_{0,0}$	$n_{0,1}$
$S \geq s$	$n_{1,0}$	$n_{1,1}$
	$n_0 = n_{0,0} + n_{1,0}$	$n_1 = n_{0,1} + n_{1,1}$

where $n_{i,j}$ represents the number of observations of the cell (i, j)

Choice of the optimal cut-off

Confusion matrix

- We notice that each cell of this table can be interpreted as follows:

	$Y = 0$	$Y = 1$
$S < s$	It is rejected and it is a bad risk (true negative)	It is rejected, but it is a good risk (false negative)
$S \geq s$	It is accepted, but it is a bad risk (false positive)	It is accepted and it is a good risk (true positive)
	(negative)	(positive)

- The cells $(S < s, Y = 0)$ and $(S \geq s, Y = 1)$ correspond to observations that are well-classified: true negative (TN) and true positive (TP)
- The cells $(S \geq s, Y = 0)$ and $(S < s, Y = 1)$ correspond to two types of errors:
  1. a false positive (FP) can induce a future loss, because it may default: this is a type I error
  2. a false negative (FN) potentially corresponds to a loss of a future P&L: this is a type II error

Choice of the optimal cut-off

Classification ratios

- We have

True Positive Rate	$\mathrm{TPR} = \dfrac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$
False Negative Rate	$\mathrm{FNR} = \dfrac{\mathrm{FN}}{\mathrm{FN} + \mathrm{TP}} = 1 - \mathrm{TPR}$
True Negative Rate	$\mathrm{TNR} = \dfrac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}$
False Positive Rate	$\mathrm{FPR} = \dfrac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} = 1 - \mathrm{TNR}$

- The true positive rate (TPR) is also known as the sensitivity or the recall
- It measures the proportion of real good risks that are correctly predicted good risk

Choice of the optimal cut-off

Classification ratios

- The precision or the positive predictive value (PPV) is

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

  It measures the proportion of predicted good risks that are correctly real good risk
- The accuracy considers the classification of both negatives and positives:

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{P} + \mathrm{N}} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{TN} + \mathrm{FP}}$$

- The F_1 score is the harmonic mean of precision and sensitivity:

$$F_1 = \frac{2}{1/\mathrm{precision} + 1/\mathrm{sensitivity}} = \frac{2 \cdot \mathrm{PPV} \cdot \mathrm{TPR}}{\mathrm{PPV} + \mathrm{TPR}}$$

Choice of the optimal cut-off

Classification ratios

Table: Confusion matrix of three scoring systems and three cut-off values s

Score	$s = 100$		$s = 200$		$s = 500$	
S_1	386	616	698	1 304	1 330	3 672
	1 614	7 384	1 302	6 696	670	4 328
S_2	372	632	700	1 304	1 386	3 616
	1 628	7 368	1 300	6 696	614	4 384
S_3	382	616	656	1 344	1 378	3 624
	1 618	7 384	1 344	6 656	622	4 376
Perfect	1 000	0	2 000	0	2 000	3 000
	1 000	8 000	0	8 000	0	5 000

Choice of the optimal cut-off

Classification ratios

Table: Binary classification ratios (in %) of the three scoring systems

Score	s	TPR	FNR	TNR	FPR	PPV	ACC	F_1
S_1	100	92.3	7.7	19.3	80.7	82.1	77.7	86.9
	200	83.7	16.3	34.9	65.1	83.7	73.9	83.7
	500	54.1	45.9	66.5	33.5	86.6	56.6	66.6
S_2	100	92.1	7.9	18.6	81.4	81.9	77.4	86.7
	200	83.7	16.3	35.0	65.0	83.7	74.0	83.7
	500	54.8	45.2	69.3	30.7	87.7	57.7	67.5
S_3	100	92.3	7.7	19.1	80.9	82.0	77.7	86.9
	200	83.2	16.8	32.8	67.2	83.2	73.1	83.2
	500	54.7	45.3	68.9	31.1	87.6	57.5	67.3
Perfect	100	100.0	0.0	50.0	50.0	88.9	90.0	94.1
	200	100.0	0.0	100.0	0.0	100.0	100.0	100.0
	500	62.5	37.5	100.0	0.0	100.0	70.0	76.9

Choice of the optimal cut-off

Classification ratios

Table: Best scoring system

Cut-off	TPR	FNR	TNR	FPR	PPV	ACC	F_1
100	S_1/S_3	S_1/S_3	S_1	S_1	S_1	S_1	S_1
200	S_1/S_2	S_1/S_2	S_2	S_2	S_2	S_2	S_2
500	S_2	S_2	S_2	S_2	S_2	S_2	S_2

Exercises

- Exercise 15.4.5 – Two-class separation maximization
- Exercise 15.4.6 – Maximum likelihood estimation of the probit model

References

GOURIÉROUX, C., and JASIAK, J. (2007)
The Econometrics of Individual Risk: Credit, Insurance, and Marketing, Princeton University Press.

RONCALLI, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapter 15.

Course 2023-2024 in Financial Risk Management
Tutorial Session 1

Thierry Roncalli*

*Amundi Asset Management[24]

*University of Paris-Saclay

September 2023

[24] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- **Tutorial Session 1: Market Risk**
- Tutorial Session 2: Credit Risk
- Tutorial Session 3: Counterparty Credit Risk and Collateral Risk
- Tutorial Session 4: Operational Risk & Asset Liability Management Risk
- Tutorial Session 5: Copulas, EVT & Stress Testing

Covariance matrix

Exercise

We consider a universe of three stocks A, B and C.

Covariance matrix

Question 1

The covariance matrix of stock returns is:

$$\Sigma = \begin{pmatrix} 4\% & & \\ 3\% & 5\% & \\ 2\% & -1\% & 6\% \end{pmatrix}$$

Covariance matrix

Question 1.a

Calculate the volatility of stock returns.

Covariance matrix

We have:

$$\sigma_A = \sqrt{\Sigma_{1,1}} = \sqrt{4\%} = 20\%$$

For the other stocks, we obtain $\sigma_B = 22.36\%$ and $\sigma_C = 24.49\%$.

Covariance matrix

Question 1.b

Deduce the correlation matrix.

Covariance matrix

The correlation is the covariance divided by the product of volatilities:

$$\rho(R_A, R_B) = \frac{\Sigma_{1,2}}{\sqrt{\Sigma_{1,1} \times \Sigma_{2,2}}} = \frac{3\%}{20\% \times 22.36\%} = 67.08\%$$

We obtain:

$$\rho = \begin{pmatrix} 100.00\% & & \\ 67.08\% & 100.00\% & \\ 40.82\% & -18.26\% & 100.00\% \end{pmatrix}$$

Covariance matrix

Question 2

We assume that the volatilities are 10%, 20% and 30%. whereas the correlation matrix is equal to:

$$\rho = \begin{pmatrix} 100\% & & \\ 50\% & 100\% & \\ 25\% & 0\% & 100\% \end{pmatrix}$$

Covariance matrix

Question 2.a

Write the covariance matrix.

Covariance matrix

Using the formula $\Sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$, it follows that:

$$\Sigma = \begin{pmatrix} 1.00\% & & \\ 1.00\% & 4.00\% & \\ 0.75\% & 0.00\% & 9.00\% \end{pmatrix}$$

Covariance matrix

Question 2.b

Calculate the volatility of the portfolio $(50\%, 50\%, 0)$.

Covariance matrix

We deduce that:

$$\begin{aligned}\sigma^2(w) &= 0.5^2 \times 1\% + 0.5^2 \times 4\% + 2 \times 0.5 \times 0.5 \times 1\% \\ &= 1.75\%\end{aligned}$$

and $\sigma(w) = 13.23\%$.

Covariance matrix

Question 2.c

Calculate the volatility of the portfolio $(60\%, -40\%, 0)$. Comment on this result.

Covariance matrix

It follows that:

$$\begin{aligned}\sigma^2(w) &= 0.6^2 \times 1\% + (-0.4)^2 \times 4\% + 2 \times 0.6 \times (-0.4) \times 1\% \\ &= 0.52\%\end{aligned}$$

and $\sigma(w) = 7.21\%$. This long/short portfolio has a lower volatility than the previous long-only portfolio, because part of the risk is hedged by the positive correlation between stocks A and B.

Covariance matrix

Question 2.d

We assume that the portfolio is long \$150 in stock A, long \$500 in stock B and short \$200 in stock C. Find the volatility of this long/short portfolio.

Covariance matrix

We have:

$$\begin{aligned}\sigma^2(w) &= 150^2 \times 1\% + 500^2 \times 4\% + (-200)^2 \times 9\% + \\ &\quad 2 \times 150 \times 500 \times 1\% + \\ &\quad 2 \times 150 \times (-200) \times 0.75\% + \\ &\quad 2 \times 500 \times (-200) \times 0\% \\ &= 14\,875\end{aligned}$$

The volatility is equal to \$121.96 and is measured in USD contrary to the two previous results which were expressed in %.

Covariance matrix

Question 3

We consider that the vector of stock returns follows a one-factor model:

$$R = \beta \mathcal{F} + \varepsilon$$

We assume that $\mathcal{F}$ and ε are independent. We note $\sigma_{\mathcal{F}}^2$ the variance of $\mathcal{F}$ and $D = \operatorname{diag}\left(\tilde{\sigma}_1^2, \tilde{\sigma}_2^2, \tilde{\sigma}_3^2\right)$ the covariance matrix of idiosyncratic risks ε_t. We use the following numerical values: $\sigma_{\mathcal{F}} = 50\%$, $\beta_1 = 0.9$, $\beta_2 = 1.3$, $\beta_3 = 0.1$, $\tilde{\sigma}_1 = 5\%$, $\tilde{\sigma}_2 = 5\%$ and $\tilde{\sigma}_3 = 15\%$.

Covariance matrix

Question 3.a

Calculate the volatility of stock returns.

Covariance matrix

We have:

$$\mathbb{E}[R] = \beta\mathbb{E}[\mathcal{F}] + \mathbb{E}[\varepsilon]$$

and:

$$R - \mathbb{E}[R] = \beta(\mathcal{F} - \mathbb{E}[\mathcal{F}]) + \varepsilon - \mathbb{E}[\varepsilon]$$

It follows that:

$$\begin{aligned} \operatorname{cov}(R) &= \mathbb{E}\left[(R - \mathbb{E}[R])(R - \mathbb{E}[R])^\top\right] \\ &= \mathbb{E}\left[\beta(\mathcal{F} - \mathbb{E}[\mathcal{F}])(\mathcal{F} - \mathbb{E}[\mathcal{F}])\beta^\top\right] + \\ &\quad 2 \times \mathbb{E}\left[\beta(\mathcal{F} - \mathbb{E}[\mathcal{F}])(\varepsilon - \mathbb{E}[\varepsilon])^\top\right] + \\ &\quad \mathbb{E}\left[(\varepsilon - \mathbb{E}[\varepsilon])(\varepsilon - \mathbb{E}[\varepsilon])^\top\right] \\ &= \sigma_{\mathcal{F}}^2\beta\beta^\top + D \end{aligned}$$

Covariance matrix

We deduce that:

$$\sigma\left(R_i\right)=\sqrt{\sigma_{\mathcal{F}}^2\beta_i^2+\tilde{\sigma}_i^2}$$

We obtain $\sigma\left(R_A\right)=18.68\%$, $\sigma\left(R_B\right)=26.48\%$ and $\sigma\left(R_C\right)=15.13\%$.

Covariance matrix

Question 3.b

Calculate the correlation between stock returns.

Covariance matrix

The correlation between stocks i and j is defined as follows:

$$\rho\left(R_i, R_j\right) = \frac{\sigma_{\mathcal{F}}^2 \beta_i \beta_j}{\sigma\left(R_i\right) \sigma\left(R_j\right)}$$

We obtain:

$$\rho = \left(\begin{array}{rrr} 100.00\% & & \\ 94.62\% & 100.00\% & \\ 12.73\% & 12.98\% & 100.00\% \end{array}\right)$$

Expected shortfall of an equity portfolio

Exercise

We consider an investment universe, which is composed of two stocks A and B. The current prices of the two stocks are respectively equal to \$100 and \$200. Their volatilities are equal to 25% and 20% whereas the cross-correlation is equal to -20%. The portfolio is long of 4 stocks A and 3 stocks B.

Expected shortfall of an equity portfolio

Question 1

Calculate the Gaussian expected shortfall at the 97.5% confidence level for a ten-day time horizon.

Expected shortfall of an equity portfolio

We have:

$$\begin{aligned}\Pi &= 4\left(P_{A,t+h} - P_{A,t}\right) + 3\left(P_{B,t+h} - P_{B,t}\right) \\ &= 4P_{A,t}R_{A,t+h} + 3P_{B,t}R_{B,t+h} \\ &= 400 \times R_{A,t+h} + 600 \times R_{B,t+h}\end{aligned}$$

where $R_{A,t+h}$ and $R_{B,t+h}$ are the stock returns for the period $[t, t+h]$. We deduce that the variance of the P&L is:

$$\begin{aligned}\sigma^2(\Pi) &= 400 \times (25\%)^2 + 600 \times (20\%)^2 + \\ & \quad 2 \times 400 \times 600 \times (-20\%) \times 25\% \times 20\% \\ &= 19\,600\end{aligned}$$

Expected shortfall of an equity portfolio

We deduce that $\sigma(\Pi) = \$140$. We know that the one-year expected shortfall is a linear function of the volatility:

$$\begin{aligned} \mathrm{ES}_\alpha(w; \text{one year}) &= \frac{\phi\left(\Phi^{-1}(\alpha)\right)}{1-\alpha} \times \sigma(\Pi) \\ &= 2.34 \times 140 \\ &= \$327.60 \end{aligned}$$

The 10-day expected shortfall is then equal to \$64.25:

$$\begin{aligned} \mathrm{ES}_\alpha(w; \text{ten days}) &= \sqrt{\frac{10}{260}} \times 327.60 \\ &= \$64.25 \end{aligned}$$

Expected shortfall of an equity portfolio

Question 2

The eight worst scenarios of daily stock returns among the last 250 historical scenarios are the following:

s	1	2	3	4	5	6	7	8
R_A	-3%	-4%	-3%	-5%	-6%	$+3\%$	$+1\%$	-1%
R_B	-4%	$+1\%$	-2%	-1%	$+2\%$	-7%	-3%	-2%

Calculate then the historical expected shortfall at the 97.5% confidence level for a ten-day time horizon.

Expected shortfall of an equity portfolio

We have:

$$\Pi_s = 400 \times R_{A,s} + 600 \times R_{B,s}$$

We deduce that the value Π_s of the daily P&L for each scenario s is:

s	1	2	3	4	5	6	7	8
Π_s	-36	-10	-24	-26	-12	-30	-14	-16
$\Pi_{s:250}$	-36	-30	-26	-24	-16	-14	-12	-10

Expected shortfall of an equity portfolio

The value-at-risk at the 97.5% confidence level correspond to the 6.25^{th} order statistic[25]. We deduce that the historical expected shortfall for a one-day time horizon is equal to:

$$\begin{aligned}\mathrm{ES}_\alpha(w;\text{one day}) &= -\mathbb{E}\left[\Pi \mid \Pi \leq -\mathrm{VaR}_\alpha(\Pi)\right] \\ &= -\frac{1}{6}\sum_{s=1}^{6}\Pi_{s:250} \\ &= \frac{1}{6}(36+30+26+24+16+14) \\ &= 24.33\end{aligned}$$

By considering the square-root-of-time rule, it follows that the 10-day expected shortfall is equal to \$76.95.

[25] We have $2.5\% \times 250 = 6.25$.

Value-at-risk of a long/short portfolio

Exercise

We consider a long/short portfolio composed of a long (buying) position in asset A and a short (selling) position in asset B. The long exposure is \$2 mn whereas the short exposure is \$1 mn. Using the historical prices of the last 250 trading days of assets A and B, we estimate that the asset volatilities σ_A and σ_B are both equal to 20% per year and that the correlation $\rho_{A,B}$ between asset returns is equal to 50%. In what follows, we ignore the mean effect.

Value-at-risk of a long/short portfolio

We note $S_{A,t}$ (resp. $S_{B,t}$) the price of stock A (resp. B) at time t. The portfolio value is:

$$P_t(w) = w_A S_{A,t} + w_B S_{B,t}$$

where w_A and w_B are the number of stocks A and B. We deduce that the P&L between t and $t+1$ is:

$$\begin{aligned}
\Pi(w) &= P_{t+1} - P_t \\
&= w_A\left(S_{A,t+1} - S_{A,t}\right) + w_B\left(S_{B,t+1} - S_{B,t}\right) \\
&= w_A S_{A,t} R_{A,t+1} + w_B S_{B,t} R_{B,t+1} \\
&= W_{A,t} R_{A,t+1} + W_{B,t} R_{B,t+1}
\end{aligned}$$

where $R_{A,t+1}$ and $R_{B,t+1}$ are the asset returns of A and B between t and $t+1$, and $W_{A,t}$ and $W_{B,t}$ are the nominal wealth invested in stocks A and B at time t.

Value-at-risk of a long/short portfolio

Question 1

Calculate the Gaussian VaR of the long/short portfolio for a one-day holding period and a 99% confidence level.

Value-at-risk of a long/short portfolio

We have $W_{A,t} = +2$ and $W_{B,t} = -1$. The P&L (expressed in USD million) has the following expression:

$$\Pi(w) = 2R_{A,t+1} - R_{B,t+1}$$

We have $\Pi(w) \sim \mathcal{N}\left(0, \sigma^2(\Pi)\right)$ with:

$$\begin{aligned}
\sigma(\Pi) &= \sqrt{(2\sigma_A)^2 + (-\sigma_B)^2 + 2\rho_{A,B} \times (2\sigma_A) \times (-\sigma_B)} \\
&= \sqrt{4 \times 0.20^2 + (-0.20)^2 - 4 \times 0.5 \times 0.20^2} \\
&= \sqrt{3} \times 20\% \\
&\simeq 34.64\%
\end{aligned}$$

Value-at-risk of a long/short portfolio

The annual volatility of the long/short portfolio is then equal to \$346 400. We consider the square-root-of-time rule to calculate the daily value-at-risk:

$$\begin{aligned}\mathrm{VaR}_{99\%}\left(w;\text{one day}\right) &= \frac{1}{\sqrt{260}}\times\Phi^{-1}\left(0.99\right)\times\sqrt{3}\times 20\% \\ &= 5.01\%\end{aligned}$$

The 99% value-at-risk is then equal to \$50 056.

Value-at-risk of a long/short portfolio

Question 2

How do you calculate the historical VaR? Using the historical returns of the last 250 trading days, the five worst scenarios of the 250 simulated daily P&L of the portfolio are $-58\,700$, $-56\,850$, $-54\,270$, $-52\,170$ and $-49\,231$. Calculate the historical VaR for a one-day holding period and a 99% confidence level.

Value-at-risk of a long/short portfolio

We use the historical data to calculate the scenarios of asset returns $(R_{A,t+1}, R_{B,t+1})$. We then deduce the empirical distribution of the P&L with the formula $\Pi(w) = 2R_{A,t+1} - R_{B,t+1}$. Finally, we calculate the empirical quantile. With 250 scenarios, the 1% decile is between the second and third worst cases:

$$\begin{aligned}\mathrm{VaR}_{99\%}(w;\text{one day}) &= -\left[-56\,850 + \frac{1}{2}(-54\,270 - (-56\,850))\right] \\ &= 55\,560\end{aligned}$$

The probability to lose \$55 560 per day is equal to 1%. We notice that the difference between the historical VaR and the Gaussian VaR is equal to 11%.

Value-at-risk of a long/short portfolio

Question 3

We assume that the multiplication factor m_c is 3. Deduce the required capital if the bank uses an internal model based on the Gaussian value-at-risk. Same question when the bank uses the historical VaR. Compare these figures with those calculated with the standardized measurement method.

Value-at-risk of a long/short portfolio

If we assume that the average of the last 60 VaRs is equal to the current VaR, we obtain:

$$\mathcal{K}^{\mathrm{IMA}} = m_c \times \sqrt{10} \times \mathrm{VaR}_{99\%}(w; \text{one day})$$

$\mathcal{K}^{\mathrm{IMA}}$ is respectively equal to \$474 877 and \$527 088 for the Gaussian and historical VaRs. In the case of the standardized measurement method, we have:

$$\begin{aligned}\mathcal{K}^{\mathrm{Specific}} &= 2 \times 8\% + 1 \times 8\% \\ &= \$240\,000\end{aligned}$$

and:

$$\begin{aligned}\mathcal{K}^{\mathrm{General}} &= |2 - 1| \times 8\% \\ &= \$80\,000\end{aligned}$$

Value-at-risk of a long/short portfolio

We deduce that:

$$\begin{aligned} \mathcal{K}^{\mathrm{SMM}} &= \mathcal{K}^{\mathrm{Specific}} + \mathcal{K}^{\mathrm{General}} \\ &= \$320\,000 \end{aligned}$$

The internal model-based approach does not achieve a reduction of the required capital with respect to the standardized measurement method. Moreover, we have to add the stressed VaR under Basel 2.5 and the IMA regulatory capital is at least multiplied by a factor of 2.

Value-at-risk of a long/short portfolio

Question 4

Show that the Gaussian VaR is multiplied by a factor equal to $\sqrt{7/3}$ if the correlation $\rho_{A,B}$ is equal to -50%. How do you explain this result?

Value-at-risk of a long/short portfolio

If $\rho_{A,B} = -0.50$, the volatility of the P&L becomes:

$$\begin{aligned} \sigma(\Pi) &= \sqrt{4 \times 0.20^2 + (-0.20)^2 - 4 \times (-0.5) \times 0.20^2} \\ &= \sqrt{7} \times 20\% \end{aligned}$$

We deduce that:

$$\frac{\mathrm{VaR}_\alpha(\rho_{A,B} = -50\%)}{\mathrm{VaR}_\alpha(\rho_{A,B} = +50\%)} = \frac{\sigma(\Pi; \rho_{A,B} = -50\%)}{\sigma(\Pi; \rho_{A,B} = +50\%)} = \sqrt{\frac{7}{3}} = 1.53$$

The value-at-risk increases because the hedging effect of the positive correlation vanishes. With a negative correlation, a long/short portfolio becomes more risky than a long-only portfolio.

Value-at-risk of a long/short portfolio

Question 5

The portfolio manager sells a call option on the stock A. The delta of the option is equal to 50%. What does the Gaussian value-at-risk of the long/short portfolio become if the nominal of the option is equal to \$2 mn? Same question when the nominal of the option is equal to \$4 mn. How do you explain this result?

Value-at-risk of a long/short portfolio

The P&L formula becomes:

$$\Pi(w) = W_{A,t} R_{A,t+1} + W_{B,t} R_{B,t+1} - \left(\mathcal{C}_{A,t+1} - \mathcal{C}_{A,t}\right)$$

where $\mathcal{C}_{A,t}$ is the call option price. We have:

$$\mathcal{C}_{A,t+1} - \mathcal{C}_{A,t} \simeq \boldsymbol{\Delta}_t \left(S_{A,t+1} - S_{A,t}\right)$$

where $\boldsymbol{\Delta}_t$ is the delta of the option. If the nominal of the option is USD 2 million, we obtain:

$$\begin{aligned} \Pi(w) &= 2R_A - R_B - 2 \times 0.5 \times R_A \\ &= R_A - R_B \end{aligned} \tag{1}$$

and:

$$\begin{aligned} \sigma(\Pi) &= \sqrt{0.20^2 + (-0.20)^2 - 2 \times 0.5 \times 0.20^2} \\ &= 20\% \end{aligned}$$

Value-at-risk of a long/short portfolio

If the nominal of the option is USD 4 million, we obtain:

$$\begin{aligned}\Pi(w) &= 2R_A - R_B - 4 \times 0.5 \times R_A \\ &= -R_B \end{aligned} \tag{2}$$

and $\sigma(\Pi) = 20\%$. In both cases, we have:

$$\begin{aligned}\mathrm{VaR}_{99\%}(w; \text{one day}) &= \frac{1}{\sqrt{260}} \times \Phi^{-1}(0.99) \times 20\% \\ &= \$28\,900\end{aligned}$$

The value-at-risk of the long/short portfolio (1) is then equal to the value-at-risk of the short portfolio (2) because of two effects: the absolute exposure of the long/short portfolio is higher than the absolute exposure of the short portfolio, but a part of the risk of the long/short portfolio is hedged by the positive correlation between the two stocks.

Value-at-risk of a long/short portfolio

Question 6

The portfolio manager replaces the short position on the stock B by selling a call option on the stock B. The delta of the option is equal to 50%. Show that the Gaussian value-at-risk is minimum when the nominal is equal to four times the correlation $\rho_{A,B}$. Deduce then an expression of the lowest Gaussian VaR. Comment on these results.

Value-at-risk of a long/short portfolio

We have:

$$\Pi(w) = W_{A,t} R_{A,t+1} - \left(\mathcal{C}_{B,t+1} - \mathcal{C}_{B,t}\right)$$

and:

$$\mathcal{C}_{B,t+1} - \mathcal{C}_{B,t} \simeq \boldsymbol{\Delta}_t \left(S_{B,t+1} - S_{B,t}\right)$$

where $\boldsymbol{\Delta}_t$ is the delta of the option. We note x the nominal of the option expressed in USD million. We obtain:

$$\begin{aligned}\Pi(w) &= 2R_A - x \times \boldsymbol{\Delta}_t \times R_B \\ &= 2R_A - \frac{x}{2} R_B\end{aligned}$$

We have[26]:

$$\begin{aligned}\sigma^2(\Pi) &= 4\sigma_A^2 + \frac{x^2\sigma_B^2}{4} + 2\rho_{A,B} \times (2\sigma_A) \times \left(-\frac{x}{2}\sigma_B\right) \\ &= \frac{\sigma_A^2}{4}\left(x^2 - 8\rho_{A,B}x + 16\right)\end{aligned}$$

[26] Because $\sigma_A = \sigma_B = 20\%$.

Value-at-risk of a long/short portfolio

Minimizing the Gaussian value-at-risk is equivalent to minimizing the variance of the P&L. We deduce that the first-order condition is:

$$\frac{\partial \sigma^2 (\Pi)}{\partial x} = \frac{\sigma_A^2}{4} (2x - 8\rho_{A,B}) = 0$$

We deduce that the minimum VaR is reached when the nominal of the option is $x = 4\rho_{A,B}$. We finally obtain:

$$\begin{aligned} \sigma (\Pi) &= \frac{\sigma_A}{2} \sqrt{16\rho_{A,B}^2 - 32\rho_{A,B}^2 + 16} \\ &= 2\sigma_A \sqrt{1 - \rho_{A,B}^2} \end{aligned}$$

and:

$$\begin{aligned} \mathrm{VaR}_{99\%} (w; \text{one day}) &= \frac{1}{\sqrt{260}} \times 2.33 \times 2 \times 20\% \times \sqrt{1 - \rho_{A,B}^2} \\ &\simeq 5.78\% \times \sqrt{1 - \rho_{A,B}^2} \end{aligned}$$

Value-at-risk of a long/short portfolio

If $\rho_{A,B}$ is negative (resp. positive), the exposure x is negative meaning that we have to buy (resp. to sell) a call option on stock B in order to hedge a part of the risk related to stock A. If $\rho_{A,B}$ is equal to zero, the exposure x is equal to zero because a position on stock B adds systematically a supplementary risk to the portfolio.

Risk management of exotic options

Exercise

Let us consider a short position on an exotic option, whose its current value $\mathcal{C}_t$ is equal to \$6.78. We assume that the price S_t of the underlying asset is \$100 and the implied volatility Σ_t is equal to 20%.

Risk management of exotic options

Let $\mathcal{C}_t$ be the option price at time t. The P&L of the trader between t and $t+1$ is:

$$\Pi = -\left(\mathcal{C}_{t+1} - \mathcal{C}_t\right)$$

The formulation of the exercise suggests that there are two main risk factors: the price of the underlying asset S_t and the implied volatility Σ_t. We then obtain:

$$\Pi = \mathcal{C}_t\left(S_t, \Sigma_t\right) - \mathcal{C}_{t+1}\left(S_{t+1}, \Sigma_{t+1}\right)$$

Risk management of exotic options

Question 1

At time $t+1$, the value of the underlying asset is \$97 and the implied volatility remains constant. We find that the P&L of the trader between t and $t+1$ is equal to \$1.37. Can we explain the P&L by the sensitivities knowing that the estimates of delta $\boldsymbol{\Delta}_t$, gamma $\boldsymbol{\Gamma}_t$ and vega[a] $\boldsymbol{\upsilon}_t$ are respectively equal to 49%, 2% and 40%?

[a] measured in volatility points.

Risk management of exotic options

We have:

$$\begin{aligned}\Pi &= C_t\left(S_t, \Sigma_t\right) - C_{t+1}\left(S_{t+1}, \Sigma_{t+1}\right) \\ &\approx -\boldsymbol{\Delta}_t\left(S_{t+1} - S_t\right) - \frac{1}{2}\boldsymbol{\Gamma}_t\left(S_{t+1} - S_t\right)^2 - \boldsymbol{\upsilon}_t\left(\Sigma_{t+1} - \Sigma_t\right)\end{aligned}$$

Using the numerical values of $\boldsymbol{\Delta}_t$, $\boldsymbol{\Gamma}_t$ and $\boldsymbol{\upsilon}_t$, we obtain:

$$\begin{aligned}\Pi &\approx -0.49 \times (97 - 100) - \frac{1}{2} \times 0.02 \times (97 - 100)^2 \\ &= 1.47 - 0.09 \\ &= 1.38\end{aligned}$$

We explain the P&L by the sensitivities very well.

Risk management of exotic options

Question 2

At time $t+2$, the price of the underlying asset is \$97 while the implied volatility increases from 20% to 22%. The value of the option $\mathcal{C}_{t+2}$ is now equal to \$6.17. Can we explain the P&L by the sensitivities knowing that the estimates of delta $\boldsymbol{\Delta}_{t+1}$, gamma $\boldsymbol{\Gamma}_{t+1}$ and vega $\boldsymbol{\upsilon}_{t+1}$ are respectively equal to 43%, 2% and 38%?

Risk management of exotic options

We have:

$$\begin{aligned}\Pi &= C_{t+1}\left(S_{t+1},\Sigma_{t+1}\right) - C_{t+2}\left(S_{t+2},\Sigma_{t+2}\right) \\ &\approx -\boldsymbol{\Delta}_{t+1}\left(S_{t+2}-S_{t+1}\right) - \frac{1}{2}\boldsymbol{\Gamma}_{t+1}\left(S_{t+2}-S_{t+1}\right)^2 - \\ &\quad \boldsymbol{\upsilon}_{t+1}\left(\Sigma_{t+2}-\Sigma_{t+1}\right)\end{aligned}$$

Using the numerical values of $\boldsymbol{\Delta}_{t+1}$, $\boldsymbol{\Gamma}_{t+1}$ and $\boldsymbol{\upsilon}_{t+1}$, we obtain:

$$\begin{aligned}\Pi &\approx -0.49\times 0 - \frac{1}{2}\times 0.02\times 0^2 - 0.38\times(22-20) \\ &= -0.76\end{aligned}$$

Risk management of exotic options

To compare this value with the true P&L, we have to calculate $\mathcal{C}_{t+1}$:

$$\begin{aligned}\mathcal{C}_{t+1} &= \mathcal{C}_t - (\mathcal{C}_t - \mathcal{C}_{t+1}) \\ &= 6.78 - 1.37 \\ &= 5.41\end{aligned}$$

We deduce that:

$$\begin{aligned}\Pi &= \mathcal{C}_{t+1} - \mathcal{C}_{t+2} \\ &= 5.41 - 6.17 \\ &= -0.76\end{aligned}$$

Again, the sensitivities explain the P&L very well.

Risk management of exotic options

Question 3

At time $t+3$, the price of the underlying asset is \$95 and the value of the implied volatility is 19%. We find that the P&L of the trader between $t+2$ and $t+3$ is equal to \$0.58. Can we explain the P&L by the sensitivities knowing that the estimates of delta $\boldsymbol{\Delta}_{t+2}$, gamma $\boldsymbol{\Gamma}_{t+2}$ and vega $\boldsymbol{\upsilon}_{t+2}$ are respectively equal to 44%, 1.8% and 38%.

Risk management of exotic options

We have:

$$\begin{aligned}\Pi &= C_{t+2}\left(S_{t+2},\Sigma_{t+2}\right)-C_{t+3}\left(S_{t+3},\Sigma_{t+3}\right) \\ &\approx -\boldsymbol{\Delta}_{t+2}\left(S_{t+3}-S_{t+2}\right)-\frac{1}{2}\boldsymbol{\Gamma}_{t+2}\left(S_{t+3}-S_{t+2}\right)^{2}- \\ &\quad \boldsymbol{\upsilon}_{t+2}\left(\Sigma_{t+3}-\Sigma_{t+2}\right)\end{aligned}$$

Using the numerical values of $\boldsymbol{\Delta}_{t+2}$, $\boldsymbol{\Gamma}_{t+2}$ and $\boldsymbol{\upsilon}_{t+2}$, we obtain:

$$\begin{aligned}\Pi &\approx -0.44\times(95-97)-\frac{1}{2}\times 0.018\times(95-97)^{2}- \\ &\quad 0.38\times(19-22) \\ &= 0.88-0.036+1.14 \\ &= 1.984\end{aligned}$$

The P&L approximated by the Greek coefficients largely overestimate the true value of the P&L.

Risk management of exotic options

Question 4

What can we conclude in terms of model risk?

Risk management of exotic options

We notice that the approximation using the Greek coefficients works very well when one risk factor remains constant:

- Between t and $t+1$, the price of the underlying asset changes, but not the implied volatility;
- Between $t+1$ and $t+2$, this is the implied volatility that changes whereas the price of the underlying asset is constant.

Therefore, we can assume that the bad approximation between $t+2$ and $t+3$ is due to the cross effect between S_t and Σ_t. In terms of model risk, the P&L is then exposed to the vanna risk, meaning that the Black-Scholes model is not appropriate to price and hedge this exotic option.

Course 2023-2024 in Financial Risk Management
Tutorial Session 2

Thierry Roncalli*

*Amundi Asset Management[27]

*University of Paris-Saclay

September 2023

[27] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Tutorial Session 1: Market Risk
- **Tutorial Session 2: Credit Risk**
- Tutorial Session 3: Counterparty Credit Risk and Collateral Risk
- Tutorial Session 4: Operational Risk & Asset Liability Management Risk
- Tutorial Session 5: Copulas, EVT & Stress Testing

Single and multi-name credit default swaps

Question 1

We assume that the default time $\boldsymbol{\tau}$ follows an exponential distribution with parameter λ. Write the cumulative distribution function $\mathbf{F}$, the survival function $\mathbf{S}$ and the density function f of the random variable $\boldsymbol{\tau}$. How do we simulate this default time?

Single and multi-name credit default swaps

We have $\mathbf{F}(t) = 1 - e^{-\lambda t}$, $\mathbf{S}(t) = e^{-\lambda t}$ and $f(t) = \lambda e^{-\lambda t}$. We know that $\mathbf{S}(\boldsymbol{\tau}) \sim \mathcal{U}_{[0,1]}$. Indeed, we have:

$$\begin{aligned} \Pr\{U \leq u\} &= \Pr\{\mathbf{S}(\boldsymbol{\tau}) \leq u\} \\ &= \Pr\left\{\boldsymbol{\tau} \geq \mathbf{S}^{-1}(u)\right\} \\ &= \mathbf{S}\left(\mathbf{S}^{-1}(u)\right) \\ &= u \end{aligned}$$

It follows that $\boldsymbol{\tau} = \mathbf{S}^{-1}(U)$ with $U \sim \mathcal{U}_{[0,1]}$. Let u be a uniform random variate. Simulating τ is then equivalent to transform u into t:

$$t = -\frac{1}{\lambda} \ln u$$

Single and multi-name credit default swaps

Question 2

We consider a CDS 3M with two-year maturity and $1 mn notional principal. The recovery rate $\mathcal{R}$ is equal to 40% whereas the spread $\mathcal{s}$ is equal to 150 bps at the inception date. We assume that the protection leg is paid at the default time.

Single and multi-name credit default swaps

Question 2.a

Give the cash flow chart. What is the P&L of the protection seller A if the reference entity does not default? What is the PnL of the protection buyer B if the reference entity defaults in one year and two months?

Single and multi-name credit default swaps

The premium leg is paid quarterly. The coupon payment is then equal to:

$$
\begin{aligned}
\mathcal{PL}(t_m) &= \Delta t_m \times \mathcal{s} \times N \\
&= \frac{1}{4} \times 150 \times 10^{-4} \times 10^{6} \\
&= \$3\,750
\end{aligned}
$$

In case of default, the default leg paid by protection seller is equal to:

$$
\begin{aligned}
\mathcal{DL} &= (1 - \mathcal{R}) \times N \\
&= (1 - 40\%) \times 10^{6} \\
&= \$600\,000
\end{aligned}
$$

Single and multi-name credit default swaps

The corresponding cash flow chart is given in Figure 189. If the reference entity does not default, the P&L of the protection seller is the sum of premium interests:

$$\Pi^{\text{seller}} = 8 \times 3\,750 = \$30\,000$$

If the reference entity defaults in one year and two months, the P&L of the protection buyer is[28]:

$$\begin{aligned}\Pi^{\text{buyer}} &= (1-\mathcal{R}) \times N - \sum_{t_m < \boldsymbol{\tau}} \Delta t_m \times \mathcal{s} \times N \\ &= (1-40\%) \times 10^6 - \left(4 + \frac{2}{3}\right) \times 3\,750 \\ &= \$582\,500\end{aligned}$$

[28] We include the accrued premium.

Single and multi-name credit default swaps

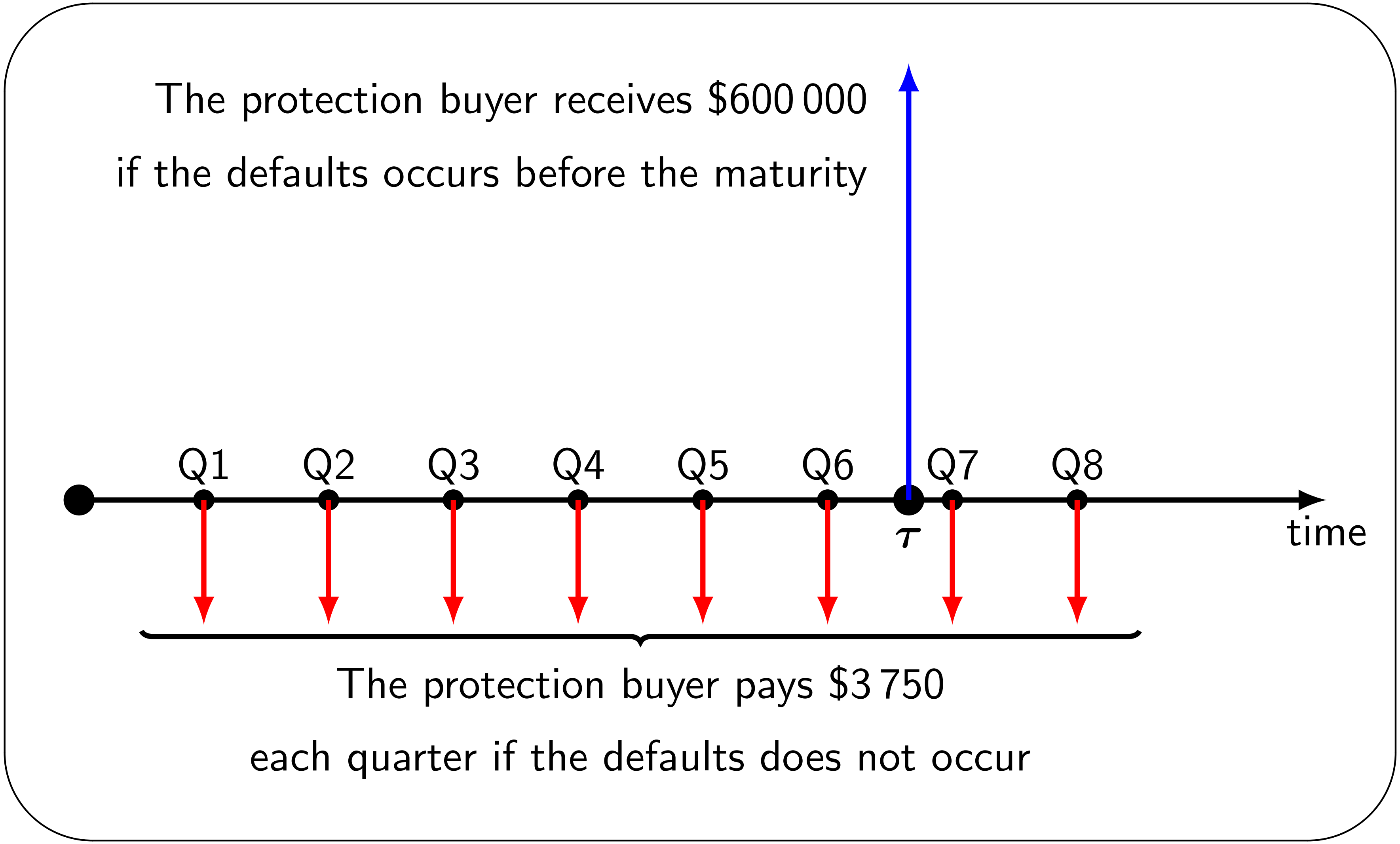


Figure: Cash flow chart of the CDS contract

Single and multi-name credit default swaps

Question 2.b

What is the relationship between $\mathcal{s}$, $\mathcal{R}$ and λ? What is the implied one-year default probability at the inception date?

Single and multi-name credit default swaps

Using the credit triangle relationship, we have:

$$s \simeq (1 - \mathcal{R}) \times \lambda$$

We deduce that[29]:

$$\begin{aligned} \mathrm{PD} &\simeq \lambda \\ &\simeq \frac{s}{1 - \mathcal{R}} \\ &= \frac{150 \times 10^{-4}}{1 - 40\%} \\ &= 2.50\% \end{aligned}$$

[29] We recall that the one-year default probability is approximately equal to λ:

$$\begin{aligned} \mathrm{PD} &= 1 - \mathbf{S}(1) \\ &= 1 - e^{-\lambda} \\ &\simeq 1 - (1 - \lambda) \\ &\simeq \lambda \end{aligned}$$

Single and multi-name credit default swaps

Question 2.c

Seven months later, the CDS spread has increased and is equal to 450 bps. Estimate the new default probability. The protection buyer B decides to realize his P&L. For that, he reassigns the CDS contract to the counterparty C. Explain the offsetting mechanism if the risky PV01 is equal to 1.189.

Single and multi-name credit default swaps

We denote by $\mathcal{s}'$ the new CDS spread. The default probability becomes:

$$\begin{aligned}\mathrm{PD} &= \frac{\mathcal{s}'}{1-\mathcal{R}} \\ &= \frac{450 \times 10^{-4}}{1-40\%} \\ &= 7.50\%\end{aligned}$$

The protection buyer is short credit and benefits from the increase of the default probability. His mark-to-market is therefore equal to:

$$\begin{aligned}\Pi^{\mathrm{buyer}} &= N \times (\mathcal{s}' - \mathcal{s}) \times \mathrm{RPV}_{01} \\ &= 10^{6} \times (450 - 150) \times 10^{-4} \times 1.189 \\ &= \$35\,671\end{aligned}$$

The offsetting mechanism is then the following: the protection buyer B transfers the agreement to C, who becomes the new protection buyer; C continues to pay a premium of 150 bps to the protection seller A; in return, C pays a cash adjustment of \$35 671 to B.

Single and multi-name credit default swaps

Question 3

We consider the following CDS spread curves for three reference entities:

Maturity	#1	#2	#3
6M	130 bps	1 280 bps	30 bps
1Y	135 bps	970 bps	35 bps
3Y	140 bps	750 bps	50 bps
5Y	150 bps	600 bps	80 bps

Single and multi-name credit default swaps

Question 3.a

Define the notion of credit curve. Comment the previous spread curves.

Single and multi-name credit default swaps

For a given date t, the credit curve is the relationship between the maturity T and the spread $\mathcal{S}_t(T)$. The credit curve of the reference entity #1 is almost flat. For the entity #2, the spread is very high in the short-term, meaning that there is a significative probability that the entity defaults. However, if the entity survive, the market anticipates that it will improve its financial position in the long-run. This explains that the credit curve #2 is decreasing. For reference entity #3, we obtain opposite conclusions. The company is actually very strong, but there are some uncertainties in the future[30]. The credit curve is then increasing.

[30] An example is a company whose has a monopoly because of a strong technology, but faces a hard competition because technology is evolving fast in its domain (e.g. Blackberry at the end of 2000s).

Single and multi-name credit default swaps

Question 3.b

Using the Merton Model, we estimate that the one-year default probability is equal to 2.5% for #1, 5% for #2 and 2% for #3 at a five-year horizon time. Which arbitrage position could we consider about the reference entity #2?

Single and multi-name credit default swaps

If we consider a standard recovery rate (40%), the implied default probability is 2.50% for #1, 10% for #2 and 1.33% for #3. We can consider a short credit position in #2. In this case, we sell the 5Y protection on #2 because the model tells us that the market default probability is over-estimated. In place of this directional bet, we could consider a relative value strategy: selling the 5Y protection on #2 and buying the 5Y protection on #3.

Single and multi-name credit default swaps

Question 4

We consider a basket of n single-name CDS.

Single and multi-name credit default swaps

Question 4.a

What is a first-to-default (FtD), a second-to-default (StD) and a last-to-default (LtD)?

Single and multi-name credit default swaps

Let $\boldsymbol{\tau}_{k:n}$ be the k^{th} default among the basket. FtD, StD and LtD are three CDS products, whose credit event is related to the default times $\boldsymbol{\tau}_{1:n}$, $\boldsymbol{\tau}_{2:n}$ and $\boldsymbol{\tau}_{n:n}$.

Single and multi-name credit default swaps

Question 4.b

Define the notion of default correlation˙What is its impact on three previous spreads?

Single and multi-name credit default swaps

The default correlation ρ measures the dependence between two default times $\boldsymbol{\tau}_i$ and $\boldsymbol{\tau}_j$. The spread of the FtD (resp. LtD) is a decreasing (resp. increasing) function with respect to ρ.

Single and multi-name credit default swaps

Question 4.c

We assume that $n = 3$. Show the following relationship:

$$s_1^{\mathrm{CDS}} + s_2^{\mathrm{CDS}} + s_3^{\mathrm{CDS}} = s^{\mathrm{FtD}} + s^{\mathrm{StD}} + s^{\mathrm{LtD}}$$

where s_i^{CDS} is the CDS spread of the i^{th} reference entity.

Single and multi-name credit default swaps

To fully hedge the credit portfolio of the 3 entities, we can buy the 3 CDS. Another solution is to buy the FtD plus the StD and the LtD (or the third-to-default). Because these two hedging strategies are equivalent, we deduce that:

$$\mathcal{s}_1^{\mathrm{CDS}} + \mathcal{s}_2^{\mathrm{CDS}} + \mathcal{s}_3^{\mathrm{CDS}} = \mathcal{s}^{\mathrm{FtD}} + \mathcal{s}^{\mathrm{StD}} + \mathcal{s}^{\mathrm{LtD}}$$

Single and multi-name credit default swaps

Question 4.d

Many professionals and academics believe that the subprime crisis is due to the use of the Normal copula. Using the results of the previous question, what could you conclude?

Single and multi-name credit default swaps

We notice that the default correlation does not affect the value of the CDS basket, but only the price distribution between FtD, StD and LtD. We obtain a similar result for CDO[31]. In the case of the subprime crisis, all the CDO tranches have suffered, meaning that the price of the underlying basket has dropped. The reasons were the underestimation of default probabilities.

[31] The junior, mezzanine and senior tranches can be viewed as FtD, StD and LtD.

Risk contribution in the Basel II model

Question 1

We note L the portfolio loss of n credit and w_i the exposure at default of the i^{th} credit. We have:

$$L(w) = w^\top \varepsilon = \sum_{i=1}^{n} w_i \times \varepsilon_i \tag{3}$$

where ε_i is the unit loss of the i^{th} credit. Let $\mathbf{F}$ be the cumulative distribution function of $L(w)$.

Risk contribution in the Basel II model

Question 1.a

We assume that $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \sim \mathcal{N}(\mathbf{0}, \Sigma)$. Compute the value-at-risk $\mathrm{VaR}_\alpha(w)$ of the portfolio when the confidence level is equal to α.

Risk contribution in the Basel II model

The portfolio loss L follows a Gaussian probability distribution:

$$L(w) \sim \mathcal{N}\left(0, \sqrt{w^\top \Sigma w}\right)$$

We deduce that:

$$\mathrm{VaR}_\alpha(w) = \Phi^{-1}(\alpha)\sqrt{w^\top \Sigma w}$$

Risk contribution in the Basel II model

Question 1.b

Deduce the marginal value-at-risk of the i^{th} credit. Define then the risk contribution $\mathcal{RC}_i$ of the i^{th} credit.

Risk contribution in the Basel II model

We have:

$$\begin{aligned}\frac{\partial \operatorname{VaR}_\alpha(w)}{\partial w} &= \frac{\partial}{\partial w}\left(\Phi^{-1}(\alpha)\left(w^\top \Sigma w\right)^{\frac{1}{2}}\right) \\ &= \Phi^{-1}(\alpha)\frac{1}{2}\left(w^\top \Sigma w\right)^{-\frac{1}{2}}(2\Sigma w) \\ &= \Phi^{-1}(\alpha)\frac{\Sigma w}{\sqrt{w^\top \Sigma w}}\end{aligned}$$

The marginal value-at-risk of the i^{th} credit is then:

$$\mathcal{MR}_i = \frac{\partial \operatorname{VaR}_\alpha(w)}{\partial w_i} = \Phi^{-1}(\alpha)\frac{(\Sigma w)_i}{\sqrt{w^\top \Sigma w}}$$

The risk contribution of the i^{th} credit is the product of the exposure by the marginal risk:

$$\begin{aligned}\mathcal{RC}_i &= w_i \times \mathcal{MR}_i \\ &= \Phi^{-1}(\alpha)\frac{w_i \times (\Sigma w)_i}{\sqrt{x^\top \Sigma x}}\end{aligned}$$

Risk contribution in the Basel II model

Question 1.c

Check that the marginal value-at-risk is equal to:

$$\frac{\partial \operatorname{VaR}_{\alpha}(w)}{\partial w_i} = \mathbb{E}\left[\varepsilon_i \mid L(w) = \mathbf{F}^{-1}(\alpha)\right]$$

Comment on this result.

Risk contribution in the Basel II model

By construction, the random vector $(\varepsilon, L(w))$ is Gaussian with:

$$\begin{pmatrix} \varepsilon \\ L(w) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma w \\ w^\top \Sigma & w^\top \Sigma w \end{pmatrix} \right)$$

We deduce that the conditional distribution function of ε given that $L(w) = \ell$ is Gaussian and we have:

$$\mathbb{E}\left[\varepsilon \mid L(w) = \ell\right] = \mathbf{0} + \Sigma w \left(w^\top \Sigma w\right)^{-1} (\ell - 0)$$

We finally obtain:

$$\begin{aligned} \mathbb{E}\left[\varepsilon \mid L(w) = \mathbf{F}^{-1}(\alpha)\right] &= \Sigma w \left(w^\top \Sigma w\right)^{-1} \Phi^{-1}(\alpha) \sqrt{w^\top \Sigma w} \\ &= \Phi^{-1}(\alpha) \frac{\Sigma w}{\sqrt{w^\top \Sigma w}} \\ &= \frac{\partial \operatorname{VaR}_\alpha(w)}{\partial w} \end{aligned}$$

The marginal VaR of the i^{th} credit is then equal to the conditional mean of the individual loss ε_i given that the portfolio loss is exactly equal to the value at risk

Risk contribution in the Basel II model

Question 2

We consider the Basel II model of credit risk and the value-at-risk risk measure. The expression of the portfolio loss is given by:

$$L = \sum_{i=1}^{n} \mathrm{EAD}_i \times \mathrm{LGD}_i \times \mathbb{1}\left\{\boldsymbol{\tau}_i < M_i\right\} \tag{4}$$

Risk contribution in the Basel II model

Question 2.a

Define the different parameters EAD_i, LGD_i, $\boldsymbol{\tau}_i$ and M_i. Show that Model (4) can be written as Model (3) by identifying w_i and ε_i.

Risk contribution in the Basel II model

EAD_i is the exposure at default, LGD_i is the loss given default, $\boldsymbol{\tau}_i$ is the default time and T_i is the maturity of the credit i. We have:

$$\left\{ \begin{array}{l} w_i = \mathrm{EAD}_i \\ \varepsilon_i = \mathrm{LGD}_i \times \mathbb{1}\left\{ \boldsymbol{\tau}_i < T_i \right\} \end{array} \right.$$

The exposure at default is not random, which is not the case of the loss given default.

Risk contribution in the Basel II model

Question 2.b

What are the necessary assumptions $(\mathcal{H})$ to obtain this result:

$$\mathbb{E}\left[\varepsilon_i \mid L = \mathbf{F}^{-1}(\alpha)\right] = \mathbb{E}\left[\mathrm{LGD}_i\right] \times \mathbb{E}\left[D_i \mid L = \mathbf{F}^{-1}(\alpha)\right]$$

with $D_i = \mathbb{1}\left\{\boldsymbol{\tau}_i < M_i\right\}$.

Risk contribution in the Basel II model

We have to make the following assumptions:

(i) the loss given default LGD_i is independent from the default time $\boldsymbol{\tau}_i$;

(ii) the portfolio is infinitely fine-grained meaning that there is no exposure concentration:

$$\frac{\mathrm{EAD}_i}{\sum_{i=1}^n \mathrm{EAD}_i} \simeq 0$$

(iii) the default times depend on a common risk factor X and the relationship is monotonic (increasing or decreasing).

In this case, we have:

$$\mathbb{E}\left[\varepsilon_i \mid L = \mathbf{F}^{-1}(\alpha)\right] = \mathbb{E}\left[\mathrm{LGD}_i\right] \times \mathbb{E}\left[D_i \mid L = \mathbf{F}^{-1}(\alpha)\right]$$

with $D_i = \mathbb{1}\left\{\boldsymbol{\tau}_i < T_i\right\}$.

Risk contribution in the Basel II model

Question 2.c

Deduce the risk contribution $\mathcal{RC}_i$ of the i^{th} credit and the value-at-risk of the credit portfolio.

Risk contribution in the Basel II model

It follows that:

$$\begin{aligned}\mathcal{RC}_i &= w_i \times \mathcal{MR}_i \\ &= \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \mathbb{E}\left[D_i \mid L = \mathbf{F}^{-1}(\alpha)\right]\end{aligned}$$

The expression of the value-at-risk is then:

$$\begin{aligned}\mathrm{VaR}_\alpha(w) &= \sum_{i=1}^{n} \mathcal{RC}_i \\ &= \sum_{i=1}^{n} \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \mathbb{E}\left[D_i \mid L = \mathbf{F}^{-1}(\alpha)\right]\end{aligned}$$

Risk contribution in the Basel II model

Question 2.d

We assume that the credit i defaults before the maturity M_i if a latent variable Z_i goes below a barrier B_i:

$$\boldsymbol{\tau}_i \leq M_i \Leftrightarrow Z_i \leq B_i$$

We consider that $Z_i = \sqrt{\rho} X + \sqrt{1-\rho}\varepsilon_i$ where Z_i, X and ε_i are three independent Gaussian variables $\mathcal{N}(0,1)$. X is the factor (or the systematic risk) and ε_i is the idiosyncratic risk.

Risk contribution in the Basel II model

Question 2.d (i)

Interpret the parameter ρ.

Risk contribution in the Basel II model

We have

$$\begin{aligned}\mathbb{E}\left[Z_i Z_j\right] &= \mathbb{E}\left[\left(\sqrt{\rho}X+\sqrt{1-\rho}\varepsilon_i\right)\left(\sqrt{\rho}X+\sqrt{1-\rho}\varepsilon_j\right)\right] \\ &= \rho\end{aligned}$$

ρ is the constant correlation between assets Z_i and Z_j.

Risk contribution in the Basel II model

Question 2.d (ii)

Calculate the unconditional default probability:

$$p_i = \Pr\{\boldsymbol{\tau}_i \leq M_i\}$$

Risk contribution in the Basel II model

We have:

$$\begin{aligned} p_i &= \Pr \left\{ \tau_i \leq T_i \right\} \\ &= \Pr \left\{ Z_i \leq B_i \right\} \\ &= \Phi \left(B_i \right) \end{aligned}$$

Risk contribution in the Basel II model

Question 2.d (iii)

Calculate the conditional default probability:

$$p_i(x) = \Pr\{\boldsymbol{\tau}_i \leq M_i \mid X = x\}$$

Risk contribution in the Basel II model

It follows that:

$$\begin{aligned}
p_i(x) &= \Pr\{Z_i \le B_i \mid X = x\} \\
&= \Pr\left\{\sqrt{\rho}X + \sqrt{1-\rho}\varepsilon_i \le B_i \mid X = x\right\} \\
&= \Pr\left\{\varepsilon_i \le \frac{B_i - \sqrt{\rho}X}{\sqrt{1-\rho}} \middle| X = x\right\} \\
&= \Phi\left(\frac{B_i - \sqrt{\rho}x}{\sqrt{1-\rho}}\right) \\
&= \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}x}{\sqrt{1-\rho}}\right)
\end{aligned}$$

Risk contribution in the Basel II model

Question 2.e

Show that, under the previous assumptions $(\mathcal{H})$, the risk contribution $\mathcal{RC}_i$ of the i^{th} credit is:

$$\mathcal{RC}_i = \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \Phi\left(\frac{\Phi^{-1}(p_i) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right) \tag{5}$$

when the risk measure is the value-at-risk.

Risk contribution in the Basel II model

Under the assumptions $(\mathcal{H})$, we know that:

$$\begin{aligned} L &= \sum_{i=1}^{n} \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times p_i(X) \\ &= \sum_{i=1}^{n} \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho} X}{\sqrt{1-\rho}}\right) \\ &= g(X) \end{aligned}$$

with $g'(x) < 0$. We deduce that:

$$\begin{aligned} \mathrm{VaR}_\alpha(w) = \mathbf{F}^{-1}(\alpha) &\Leftrightarrow \Pr\{g(X) \leq \mathrm{VaR}_\alpha(w)\} = \alpha \\ &\Leftrightarrow \Pr\left\{X \geq g^{-1}(\mathrm{VaR}_\alpha(w))\right\} = \alpha \\ &\Leftrightarrow \Pr\left\{X \leq g^{-1}(\mathrm{VaR}_\alpha(w))\right\} = 1 - \alpha \\ &\Leftrightarrow g^{-1}(\mathrm{VaR}_\alpha(w)) = \Phi^{-1}(1-\alpha) \\ &\Leftrightarrow \mathrm{VaR}_\alpha(w) = g\left(\Phi^{-1}(1-\alpha)\right) \end{aligned}$$

Risk contribution in the Basel II model

It follows that:

$$\begin{aligned}\mathrm{VaR}_\alpha(w) &= g\left(\Phi^{-1}(1-\alpha)\right) \\ &= \sum_{i=1}^n \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times p_i\left(\Phi^{-1}(1-\alpha)\right)\end{aligned}$$

The risk contribution $\mathcal{RC}_i$ of the ith credit is then:

$$\begin{aligned}\mathcal{RC}_i &= \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times p_i\left(\Phi^{-1}(1-\alpha)\right) \\ &= \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}\Phi^{-1}(1-\alpha)}{\sqrt{1-\rho}}\right) \\ &= \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times \Phi\left(\frac{\Phi^{-1}(p_i) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right)\end{aligned}$$

Risk contribution in the Basel II model

Question 3

We now assume that the risk measure is the expected shortfall:

$$\mathrm{ES}_{\alpha}(w) = \mathbb{E}\left[L \mid L \geq \mathrm{VaR}_{\alpha}(w)\right]$$

Risk contribution in the Basel II model

Question 3.a

In the case of the Basel II framework, show that we have:

$$\mathrm{ES}_{\alpha}(w) = \sum_{i=1}^{n} \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times \mathbb{E}\left[p_i(X) \mid X \leq \Phi^{-1}(1-\alpha)\right]$$

Risk contribution in the Basel II model

We note Ω the event $X \leq g^{-1}(\mathrm{VaR}_\alpha(w))$ or equivalently $X \leq \Phi^{-1}(1-\alpha)$. We have:

$$\begin{aligned}
\mathrm{ES}_\alpha(w) &= \mathbb{E}[L \mid L \geq \mathrm{VaR}_\alpha(w)] \\
&= \mathbb{E}[L \mid g(X) \geq \mathrm{VaR}_\alpha(w)] \\
&= \mathbb{E}\left[L \mid X \leq g^{-1}(\mathrm{VaR}_\alpha(w))\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times p_i(X) \mid \Omega\right] \\
&= \sum_{i=1}^{n} \mathrm{EAD}_i \times \mathbb{E}[\mathrm{LGD}_i] \times \mathbb{E}[p_i(X) \mid \Omega]
\end{aligned}$$

Risk contribution in the Basel II model

Question 3.b

By using the following result:

$$\int_{-\infty}^{c} \Phi(a+bx)\phi(x)\,\mathrm{d}x = \Phi_2\left(c, \frac{a}{\sqrt{1+b^2}}; \frac{-b}{\sqrt{1+b^2}}\right)$$

where $\Phi_2(x, y; \rho)$ is the cdf of the bivariate Gaussian distribution with correlation ρ on the space $[-\infty, x] \times [-\infty, y]$, deduce that the risk contribution $\mathcal{RC}_i$ of the i^{th} credit in the Basel II model is:

$$\mathcal{RC}_i = \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \frac{\mathbf{C}\left(1-\alpha, p_i; \sqrt{\rho}\right)}{1-\alpha} \tag{6}$$

when the risk measure is the expected shortfall. Here $\mathbf{C}(u_1, u_2; \theta)$ is the Normal copula with parameter θ.

Risk contribution in the Basel II model

It follows that:

$$\begin{aligned}
\mathbb{E}\left[p_i(X) \mid \Omega\right] &= \mathbb{E}\left[\Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho}X}{\sqrt{1-\rho}}\right) \middle| \Omega\right] \\
&= \int_{-\infty}^{\Phi^{-1}(1-\alpha)} \Phi\left(\frac{\Phi^{-1}(p_i)}{\sqrt{1-\rho}} + \frac{-\sqrt{\rho}}{\sqrt{1-\rho}}x\right) \times \\
&\qquad \frac{\phi(x)}{\Phi\left(\Phi^{-1}(1-\alpha)\right)}\,\mathrm{d}x \\
&= \frac{\Phi_2\left(\Phi^{-1}(1-\alpha), \Phi^{-1}(p_i); \sqrt{\rho}\right)}{1-\alpha} \\
&= \frac{\mathbf{C}\left(1-\alpha, p_i; \sqrt{\rho}\right)}{1-\alpha}
\end{aligned}$$

where $\mathbf{C}$ is the Gaussian copula. We deduce that:

$$\mathcal{RC}_i = \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \frac{\mathbf{C}\left(1-\alpha, p_i; \sqrt{\rho}\right)}{1-\alpha}$$

Risk contribution in the Basel II model

Question 3.c

What become the results (5) and (6) if the correlation ρ is equal to zero? Same question if $\rho = 1$.

Risk contribution in the Basel II model

If $\rho = 0$, we have:

$$\begin{aligned}\Phi\left(\frac{\Phi^{-1}(p_i)+\sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right) &= \Phi\left(\Phi^{-1}(p_i)\right) \\ &= p_i\end{aligned}$$

and:

$$\begin{aligned}\frac{\mathbf{C}\left(1-\alpha, p_i; \sqrt{\rho}\right)}{1-\alpha} &= \frac{(1-\alpha)\, p_i}{1-\alpha} \\ &= p_i\end{aligned}$$

The risk contribution is the same for the value-at-risk and the expected shortfall:

$$\begin{aligned}\mathcal{RC}_i &= \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times p_i \\ &= \mathbb{E}\left[L_i\right]\end{aligned}$$

It corresponds to the expected loss of the credit.

Risk contribution in the Basel II model

If $\rho = 1$ and $\alpha > 50\%$, we have:

$$\begin{aligned} \Phi\left(\frac{\Phi^{-1}(p_i) + \sqrt{\rho}\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right) &= \lim_{\rho\to 1} \Phi\left(\frac{\Phi^{-1}(p_i) + \Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right) \\ &= 1 \end{aligned}$$

If $\rho = 1$ and α is high ($\alpha > 1 - \sup_i p_i$), we have:

$$\begin{aligned} \frac{\mathbf{C}\left(1-\alpha, p_i; \sqrt{\rho}\right)}{1-\alpha} &= \frac{\min\left(1-\alpha; p_i\right)}{1-\alpha} \\ &= 1 \end{aligned}$$

In this case, the risk contribution is the same for the value-at-risk and the expected shortfall:

$$\mathcal{RC}_i = \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right]$$

However, it does not depend on the unconditional probability of default p_i.

Risk contribution in the Basel II model

Question 4

The risk contributions (5) and (6) were obtained considering the assumptions ($\mathcal{H}$) and the default model defined in Question 2(d). What are the implications in terms of Pillar 2?

Risk contribution in the Basel II model

Pillar 2 concerns the non-compliance of assumptions $(\mathcal{H})$. In particular, we have to understand the impact on the credit risk measure if the portfolio is not infinitely fine-grained or if asset correlations are not constant.

Modeling loss given default

Question 1

What is the difference between the recovery rate and the loss given default?

Modeling loss given default

The loss given default is equal to:

$$\mathrm{LGD} = 1 - \boldsymbol{\mathcal{R}} + c$$

where c is the recovery (or litigation) cost. Consider for example a \$200 credit and suppose that the borrower defaults. If we recover \$140 and the litigation cost is \$20, we obtain $\boldsymbol{\mathcal{R}} = 70\%$ and $\mathrm{LGD} = 40\%$, but not $\mathrm{LGD} = 30\%$.

Modeling loss given default

Question 2

We consider a bank that grants 250 000 credits per year. The average amount of a credit is equal to \$50 000. We estimate that the average default probability is equal to 1% and the average recovery rate is equal to 65%. The total annual cost of the litigation department is equal to \$12.5 mn. Give an estimation of the loss given default?

Modeling loss given default

The amounts outstanding of credit is:

$$\begin{aligned} \mathrm{EAD} &= 250\,000 \times 50\,000 \\ &= \$12.5 \text{ bn} \end{aligned}$$

The annual loss after recovery is equal to:

$$\begin{aligned} L &= \mathrm{EAD} \times (1 - \mathcal{R}) \times \mathrm{PD} + C \\ &= 43.75 + 12.5 \\ &= \$56.25 \text{ mn} \end{aligned}$$

where C is the litigation cost.

Modeling loss given default

We deduce that:

$$\begin{aligned}
\mathrm{LGD} &= \frac{L}{\mathrm{EAD} \times \mathrm{PD}} \\
&= \frac{54}{12.5 \times 10^3 \times 1\%} \\
&= 45\%
\end{aligned}$$

This figure is larger than 35%, which is the loss given default without taking into account the recovery cost.

Modeling loss given default

Question 3

The probability density function of the beta probability distribution $\mathcal{B}(a,b)$ is:

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{\mathbf{B}(a,b)}$$

where $\mathbf{B}(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1}\,\mathrm{d}u$.

Modeling loss given default

Question 3.a

Why is the beta probability distribution a good candidate to model the loss given default? Which parameter pair (a, b) correspond to the uniform probability distribution?

Modeling loss given default

The Beta distribution allows to obtain all the forms of LGD (bell curve, inverted-U shaped curve, etc.). The uniform distribution corresponds to the case $\alpha = 1$ and $\beta = 1$. Indeed, we have:

$$\begin{aligned} f(x) &= \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,\mathrm{d}u} \\ &= 1 \end{aligned}$$

Modeling loss given default

Question 3.b

Let us consider a sample $(x_1, \ldots, x_n)$ of n losses in case of default. Write the log-likelihood function. Deduce the first-order conditions of the maximum likelihood estimator.

Modeling loss given default

We have:

$$\begin{aligned}\boldsymbol{\ell}(\alpha,\beta) &= \sum_{i=1}^{n} \ln f(x_i) \\ &= -n \ln \mathbf{B}(\alpha,\beta) + (\alpha-1)\sum_{i=1}^{n} \ln x_i + (\beta-1)\sum_{i=1}^{n} \ln(1-x_i)\end{aligned}$$

The first-order conditions are:

$$\frac{\partial\,\boldsymbol{\ell}(\alpha,\beta)}{\partial\,\alpha} = -n\frac{\partial_\alpha \mathbf{B}(\alpha,\beta)}{\mathbf{B}(\alpha,\beta)} + \sum_{i=1}^{n} \ln x_i = 0$$

and:

$$\frac{\partial\,\boldsymbol{\ell}(\alpha,\beta)}{\partial\,\beta} = -n\frac{\partial_\beta \mathbf{B}(\alpha,\beta)}{\mathbf{B}(\alpha,\beta)} + \sum_{i=1}^{n} \ln(1-x_i) = 0$$

Modeling loss given default

Question 3.c

We recall that the first two moments of the beta probability distribution are:

$$\begin{aligned}\mathbb{E}[X] &= \frac{a}{a+b} \\ \sigma^2(X) &= \frac{ab}{(a+b)^2(a+b+1)}\end{aligned}$$

Find the method of moments estimator.

Modeling loss given default

Let μ_{LGD} and σ_{LGD} be the mean and standard deviation of the LGD parameter. The method of moments consists in estimating α and β such that:

$$\frac{\alpha}{\alpha+\beta} = \mu_{\mathrm{LGD}}$$

and:

$$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \sigma^2_{\mathrm{LGD}}$$

We have:

$$\beta = \alpha\frac{(1-\mu_{\mathrm{LGD}})}{\mu_{\mathrm{LGD}}}$$

and:

$$(\alpha+\beta)^2(\alpha+\beta+1)\,\sigma^2_{\mathrm{LGD}} = \alpha\beta$$

Modeling loss given default

It follows that:

$$\begin{aligned}(\alpha+\beta)^2 &= \left(\alpha+\alpha\frac{(1-\mu_{\mathrm{LGD}})}{\mu_{\mathrm{LGD}}}\right)^2 \\ &= \frac{\alpha^2}{\mu_{\mathrm{LGD}}^2}\end{aligned}$$

and:

$$\alpha\beta=\frac{\alpha^2}{\mu_{\mathrm{LGD}}^2}\left(\alpha+\alpha\frac{(1-\mu_{\mathrm{LGD}})}{\mu_{\mathrm{LGD}}}+1\right)\sigma_{\mathrm{LGD}}^2=\alpha^2\frac{(1-\mu_{\mathrm{LGD}})}{\mu_{\mathrm{LGD}}}$$

We deduce that:

$$\alpha\left(1+\frac{(1-\mu_{\mathrm{LGD}})}{\mu_{\mathrm{LGD}}}\right)=\frac{(1-\mu_{\mathrm{LGD}})\,\mu_{\mathrm{LGD}}}{\sigma_{\mathrm{LGD}}^2}-1$$

Modeling loss given default

We finally obtain:

$$\hat{\alpha}_{\mathrm{MM}} = \frac{\mu_{\mathrm{LGD}}^2 \left(1 - \mu_{\mathrm{LGD}}\right)}{\sigma_{\mathrm{LGD}}^2} - \mu_{\mathrm{LGD}} \tag{7}$$

$$\hat{\beta}_{\mathrm{MM}} = \frac{\mu_{\mathrm{LGD}} \left(1 - \mu_{\mathrm{LGD}}\right)^2}{\sigma_{\mathrm{LGD}}^2} - \left(1 - \mu_{\mathrm{LGD}}\right) \tag{8}$$

Modeling loss given default

Question 4

We consider a risk class $\mathcal{C}$ corresponding to a customer/product segmentation specific to retail banking. A statistical analysis of 1 000 loss data available for this risk class gives the following results:

LGD_k	0%	25%	50%	75%	100%
n_k	100	100	600	100	100

where n_k is the number of data corresponding to LGD_k.

Modeling loss given default

Question 4.a

We consider a portfolio of 100 homogeneous credits, which belong to the risk class $\mathcal{C}$. The notional is \$10 000 whereas the annual default probability is equal to 1%. Calculate the expected loss of this credit portfolio with a one-year horizon time if we use the previous empirical distribution to model the LGD parameter.

Modeling loss given default

The mean of the loss given default is equal to:

$$\begin{aligned}\mu_{\mathrm{LGD}} &= \frac{100 \times 0\% + 100 \times 25\% + 600 \times 50\% + \ldots}{1000} \\ &= 50\%\end{aligned}$$

The expression of the expected loss is:

$$\mathrm{EL} = \sum_{i=1}^{100} \mathrm{EAD}_i \times \mathbb{E}\left[\mathrm{LGD}_i\right] \times \mathrm{PD}_i$$

where PD_i is the default probability of credit i. We finally obtain:

$$\begin{aligned}\mathrm{EL} &= \sum_{i=1}^{100} 10\,000 \times 50\% \times 1\% \\ &= \$5\,000\end{aligned}$$

Modeling loss given default

Question 4.b

We assume that the LGD parameter follows a beta distribution $\mathcal{B}(a, b)$. Calibrate the parameters a and b with the method of moments.

Modeling loss given default

We have $\mu_{\mathrm{LGD}} = 50\%$ and:

$$\begin{aligned}
\sigma_{\mathrm{LGD}} &= \sqrt{\frac{100 \times (0 - 0.5)^2 + 100 \times (0.25 - 0.5)^2 + \ldots}{1000}} \\
&= \sqrt{\frac{2 \times 0.5^2 + 2 \times 0.25^2}{10}} \\
&= \sqrt{\frac{0.625}{10}} \\
&= 25\%
\end{aligned}$$

Using Equations (7) and (8), we deduce that:

$$\begin{aligned}
\hat{\alpha}_{\mathrm{MM}} &= \frac{0.5^2 \times (1 - 0.5)}{0.25^2} - 0.5 = 1.5 \\
\hat{\beta}_{\mathrm{MM}} &= \frac{0.5 \times (1 - 0.5)^2}{0.25^2} - (1 - 0.5) = 1.5
\end{aligned}$$

Modeling loss given default

Question 4.c

We assume that the Basel II model is valid. We consider the portfolio described in Question 4(a) and calculate the unexpected loss. What is the impact if we use a uniform probability distribution instead of the calibrated beta probability distribution? Why does this result hold even if we consider different factors to model the default time?

Modeling loss given default

The previous portfolio is homogeneous and infinitely fine-grained. In this case, we know that the unexpected loss depends on the mean of the loss given default and not on the entire probability distribution. Because the expected value of the calibrated Beta distribution is 50%, there is no difference with the uniform distribution, which has also a mean equal to 50%. This result holds for the Basel model with one factor, and remains true when they are more factors.

Course 2023-2024 in Financial Risk Management
Tutorial Session 3

Thierry Roncalli*

*Amundi Asset Management[32]

*University of Paris-Saclay

September 2023

[32] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

Impact of netting agreements in counterparty credit risk

Question 1

The table below gives the current mark-to-market of 7 OTC contracts between Bank A and Bank B:

	Equity			Fixed income		FX	
	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$	$\mathcal{C}_6$	$\mathcal{C}_7$
A	$+10$	-5	$+6$	$+17$	-5	-5	$+1$
B	-11	$+6$	-3	-12	$+9$	$+5$	$+1$

The table should be read as follows: Bank A has a mark-to-market equal to 10 for the contract $\mathcal{C}_1$ whereas Bank B has a mark-to-market equal to -11 for the same contract, Bank A has a mark-to-market equal to -5 for the contract $\mathcal{C}_2$ whereas Bank B has a mark-to-market equal to $+6$ for the same contract, etc.

Impact of netting agreements in counterparty credit risk

Question 1.a

Explain why there are differences between the MtM values of a same OTC contract.

Impact of netting agreements in counterparty credit risk

Let $\mathrm{MtM}_A(\mathcal{C})$ and $\mathrm{MTM}_B(\mathcal{C})$ be the MtM values of Bank A and Bank B for the contract $\mathcal{C}$. We must theoretically verify that:

$$\begin{aligned}\mathrm{MtM}_{A+B}(\mathcal{C}) &= \mathrm{MTM}_A(\mathcal{C}) + \mathrm{MTM}_B(\mathcal{C}) \\ &= 0 \end{aligned} \tag{9}$$

In the case of listed products, the previous relationship is verified. In the case of OTC products, there are no market prices, forcing the bank to use pricing models for the valuation. The MTM value is then a mark-to-model price. Because the two banks do not use the same model with the same parameters, we note a discrepancy between the two mark-to-market prices:

$$\mathrm{MTM}_A(\mathcal{C}) + \mathrm{MTM}_B(\mathcal{C}) \neq 0$$

Impact of netting agreements in counterparty credit risk

For instance, we obtain:

$$\begin{aligned}
\mathrm{MTM}_{A+B}(\mathcal{C}_1) &= 10 - 11 = -1 \\
\mathrm{MTM}_{A+B}(\mathcal{C}_2) &= -5 + 6 = 1 \\
\mathrm{MTM}_{A+B}(\mathcal{C}_3) &= 6 - 3 = 3 \\
\mathrm{MTM}_{A+B}(\mathcal{C}_4) &= 17 - 12 = 5 \\
\mathrm{MTM}_{A+B}(\mathcal{C}_5) &= -5 + 9 = 4 \\
\mathrm{MTM}_{A+B}(\mathcal{C}_6) &= -5 + 5 = 0 \\
\mathrm{MTM}_{A+B}(\mathcal{C}_7) &= 1 + 1 = 2
\end{aligned}$$

Only the contract $\mathcal{C}_6$ satisfies the relationship (9).

Impact of netting agreements in counterparty credit risk

Question 1.b

Calculate the exposure at default of Bank A.

Impact of netting agreements in counterparty credit risk

We have:

$$\mathrm{EAD} = \sum_{i=1}^{7} \max \left(\mathrm{MTM}\left(\mathcal{C}_i\right), 0\right)$$

We deduce that:

$$\begin{aligned}
\mathrm{EAD}_A &= 10 + 6 + 17 + 1 = 34 \\
\mathrm{EAD}_B &= 6 + 9 + 5 + 1 = 21
\end{aligned}$$

Impact of netting agreements in counterparty credit risk

Question 1.c

Same question if there is a global netting agreement.

Impact of netting agreements in counterparty credit risk

If there is a global netting agreement, the exposure at default becomes:

$$\mathrm{EAD} = \max\left(\sum_{i=1}^{7} \mathrm{MTM}\left(\mathcal{C}_i\right), 0\right)$$

Using the numerical values, we obtain:

$$\begin{aligned}\mathrm{EAD}_A &= \max\left(10 - 5 + 6 + 17 - 5 - 5 + 1, 0\right) \\ &= \max\left(19, 0\right) \\ &= 19\end{aligned}$$

and:

$$\begin{aligned}\mathrm{EAD}_B &= \max\left(-11 + 6 - 3 - 12 + 9 + 5 + 1, 0\right) \\ &= \max\left(-5, 0\right) \\ &= 0\end{aligned}$$

Impact of netting agreements in counterparty credit risk

Question 1.d

Same question if the netting agreement only concerns equity products.

Impact of netting agreements in counterparty credit risk

If the netting agreement only concerns equity contracts, we have:

$$\mathrm{EAD} = \max\left(\sum_{i=1}^{3} \mathrm{MTM}\left(\mathcal{C}_i\right), 0\right) + \sum_{i=4}^{7} \max\left(\mathrm{MTM}\left(\mathcal{C}_i\right), 0\right)$$

It follows that:

$$\begin{aligned}
\mathrm{EAD}_A &= \max(10 - 5 + 6, 0) + 17 + 1 = 29 \\
\mathrm{EAD}_B &= \max(-11 + 6 - 3, 0) + 9 + 5 + 1 = 15
\end{aligned}$$

Impact of netting agreements in counterparty credit risk

Question 2

In the following, we measure the impact of netting agreements on the exposure at default.

Impact of netting agreements in counterparty credit risk

Question 2.a

We consider a first OTC contract $\mathcal{C}_1$ between Bank A and Bank B. The mark-to-market $\mathrm{MtM}_1(t)$ of Bank A for the contract $\mathcal{C}_1$ is defined as follows:

$$\mathrm{MtM}_1(t) = x_1 + \sigma_1 W_1(t)$$

where $W_1(t)$ is a Brownian motion. Calculate the potential future exposure of Bank A.

Impact of netting agreements in counterparty credit risk

The potential future exposure $e_1(t)$ is defined as follows:

$$e_1(t) = \max\left(x_1 + \sigma_1 W_1(t), 0\right)$$

We deduce that:

$$\begin{aligned} \mathbb{E}\left[e_1(t)\right] &= \int_{-\infty}^{\infty} \max(x, 0)\, f(x)\, \mathrm{d}x \\ &= \int_0^{\infty} x f(x)\, \mathrm{d}x \end{aligned}$$

where $f(x)$ is the density function of $\mathrm{MtM}_1(t)$. As we have $\mathrm{MtM}_1(t) \sim \mathcal{N}\left(x_1, \sigma_1^2 t\right)$, we deduce that:

$$\mathbb{E}\left[e_1(t)\right] = \int_0^{\infty} \frac{x}{\sigma_1\sqrt{2\pi t}} \exp\left(-\frac{1}{2}\left(\frac{x - x_1}{\sigma_1\sqrt{t}}\right)^2\right) \mathrm{d}x$$

Impact of netting agreements in counterparty credit risk

With the change of variable $y = \sigma_1^{-1} t^{-1/2} (x - x_1)$, we obtain:

$$\begin{aligned}
\mathbb{E}\left[e_1(t)\right] &= \int_{\frac{-x_1}{\sigma_1\sqrt{t}}}^{\infty} \frac{x_1 + \sigma_1\sqrt{t}y}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) \mathrm{d}y \\
&= x_1 \int_{\frac{-x_1}{\sigma_1\sqrt{t}}}^{\infty} \phi(y) \, \mathrm{d}y + \sigma_1\sqrt{t} \int_{\frac{-x_1}{\sigma_1\sqrt{t}}}^{\infty} y\phi(y) \, \mathrm{d}y \\
&= x_1 \Phi\left(\frac{x_1}{\sigma_1\sqrt{t}}\right) + \sigma_1\sqrt{t}\left[-\phi(y)\right]_{\frac{-x_1}{\sigma_1\sqrt{t}}}^{\infty} \\
&= x_1 \Phi\left(\frac{x_1}{\sigma_1\sqrt{t}}\right) + \sigma_1\sqrt{t}\phi\left(\frac{x_1}{\sigma_1\sqrt{t}}\right)
\end{aligned}$$

because $\phi(-x) = \phi(x)$ and $\Phi(-x) = 1 - \Phi(x)$.

Impact of netting agreements in counterparty credit risk

Question 2.b

We consider a second OTC contract between Bank A and Bank B. The mark-to-market is also given by the following expression:

$$\mathrm{MtM}_2(t) = x_2 + \sigma_2 W_2(t)$$

where $W_2(t)$ is a second Brownian motion that is correlated with $W_1(t)$. Let ρ be this correlation such that $\mathbb{E}[W_1(t) W_2(t)] = \rho t$. Calculate the expected exposure of bank A if there is no netting agreement.

Impact of netting agreements in counterparty credit risk

When there is no netting agreement, we have:

$$e(t) = e_1(t) + e_2(t)$$

We deduce that:

$$\begin{aligned}\mathbb{E}[e(t)] &= \mathbb{E}[e_1(t)] + \mathbb{E}[e_2(t)] \\ &= x_1 \Phi\left(\frac{x_1}{\sigma_1 \sqrt{t}}\right) + \sigma_1 \sqrt{t} \phi\left(\frac{x_1}{\sigma_1 \sqrt{t}}\right) + \\ &\quad x_2 \Phi\left(\frac{x_2}{\sigma_2 \sqrt{t}}\right) + \sigma_2 \sqrt{t} \phi\left(\frac{x_2}{\sigma_2 \sqrt{t}}\right)\end{aligned}$$

Impact of netting agreements in counterparty credit risk

Question 2.c

Same question when there is a global netting agreement between Bank A and Bank B.

Impact of netting agreements in counterparty credit risk

In the case of a netting agreement, the potential future exposure becomes:

$$\begin{aligned} e(t) &= \max\left(\mathrm{MtM}_1(t) + \mathrm{MtM}_2(t), 0\right) \\ &= \max\left(\mathrm{MtM}_{1+2}(t), 0\right) \\ &= \max\left(x_1 + x_2 + \sigma_1 W_1(t) + \sigma_2 W_2(t), 0\right) \end{aligned}$$

We deduce that:

$$\mathrm{MtM}_{1+2}(t) \sim \mathcal{N}\left(x_1 + x_2, \left(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2\right) t\right)$$

Using results of Question 2(a), we finally obtain:

$$\begin{aligned} \mathbb{E}\left[e(t)\right] &= (x_1 + x_2)\,\Phi\left(\frac{x_1 + x_2}{\sqrt{\left(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2\right) t}}\right) + \\ & \quad \sqrt{\left(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2\right) t}\,\phi\left(\frac{x_1 + x_2}{\sqrt{\left(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2\right) t}}\right) \end{aligned}$$

Impact of netting agreements in counterparty credit risk

Question 2.d

Comment on these results.

Impact of netting agreements in counterparty credit risk

We have represented the expected exposure $\mathbb{E}[e(t)]$ in Figure 190 when $x_1 = x_2 = 0$ and $\sigma_1 = \sigma_2$. We note that it is an increasing function of the time t and the volatility σ. We also observe that the netting agreement may have a big impact, especially when the correlation is low or negative.

Impact of netting agreements in counterparty credit risk

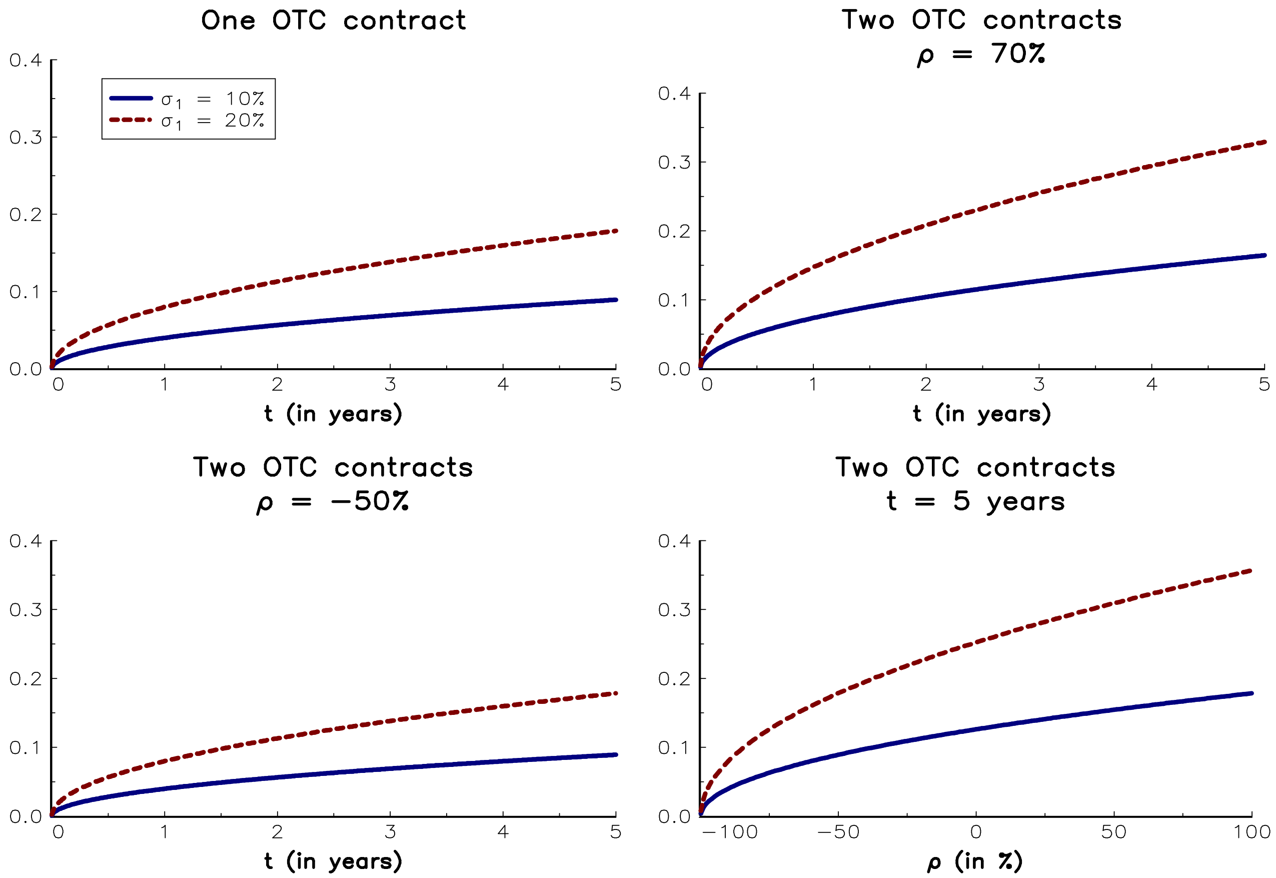


Figure: Expected exposure $\mathbb{E}[e(t)]$ when there is a netting agreement

Calculation of the CCR capital charge

We denote by $e(t)$ the potential future exposure of an OTC contract with maturity T. The current date is set to $t = 0$. Let N and σ be the notional and the volatility of the underlying contract. We assume that $e(t) = N\sigma\sqrt{t}X$ with $0 \leq X \leq 1$, $\Pr\{X \leq x\} = x^{\gamma}$ and $\gamma > 0$.

Calculation of the CCR capital charge

Question 1

Calculate the peak exposure $\mathrm{PE}_{\alpha}(t)$, the expected exposure $\mathrm{EE}(t)$ and the effective expected positive exposure $\mathrm{EEPE}(0;t)$.

Calculation of the CCR capital charge

We have:

$$\begin{aligned}
\mathbf{F}_{[0,t]}(x) &= \Pr\{e(t) \leq x\} \\
&= \Pr\left\{N\sigma\sqrt{t}U \leq x\right\} \\
&= \Pr\left\{U \leq \frac{x}{N\sigma\sqrt{t}}\right\} \\
&= \left(\frac{x}{N\sigma\sqrt{t}}\right)^{\gamma}
\end{aligned}$$

with $x \in \left[0, N\sigma\sqrt{t}\right]$. We deduce that:

$$\begin{aligned}
\mathrm{PE}_{\alpha}(t) &= \mathbf{F}_{[0,t]}^{-1}(\alpha) \\
&= N\sigma\sqrt{t}\alpha^{1/\gamma}
\end{aligned}$$

Calculation of the CCR capital charge

For the expected exposure, we obtain:

$$\begin{aligned}
\mathrm{EE}(t) &= \mathbb{E}\left[e(t)\right] \\
&= \int_0^{N\sigma\sqrt{t}} x \frac{\gamma}{\left(N\sigma\sqrt{t}\right)^{\gamma}} x^{\gamma-1}\,\mathrm{d}x \\
&= \frac{\gamma}{\left(N\sigma\sqrt{t}\right)^{\gamma}} \left[\frac{x^{\gamma+1}}{\gamma+1}\right]_0^{N\sigma\sqrt{t}} \\
&= \frac{\gamma}{\gamma+1} N\sigma\sqrt{t}
\end{aligned}$$

Calculation of the CCR capital charge

We deduce that:

$$\mathrm{EEE}(t) = \frac{\gamma}{\gamma+1} N \sigma \sqrt{t}$$

and:

$$\begin{aligned}
\mathrm{EEPE}(0;t) &= \frac{1}{t}\int_0^t \mathrm{EEE}(s)\,\mathrm{d}s \\
&= \frac{1}{t}\int_0^t \frac{\gamma}{\gamma+1} N\sigma\sqrt{s}\,\mathrm{d}s \\
&= \frac{\gamma}{\gamma+1} N\sigma \frac{1}{t}\left[\frac{2}{3}s^{3/2}\right]_0^t \\
&= \frac{2\gamma}{3(\gamma+1)} N\sigma\sqrt{t}
\end{aligned}$$

Calculation of the CCR capital charge

Question 2

The bank manages the credit risk with the foundation IRB approach and the counterparty credit risk with an internal model. We consider an OTC contract with the following parameters: N is equal to \$3 mn, the maturity T is one year, the volatility σ is set to 20% and γ is estimated at 2.

Calculation of the CCR capital charge

Question 2.a

Calculate the exposure at default EAD knowing that the bank uses the regulatory value for the parameter α.

Calculation of the CCR capital charge

When the bank uses an internal model, the regulatory exposure at default is:

$$\mathrm{EAD} = \alpha \times \mathrm{EEPE}\,(0;1)$$

Using the standard value $\alpha = 1.4$, we obtain:

$$\begin{aligned}\mathrm{EAD} &= 1.4 \times \frac{4}{9} \times 3 \times 10^{6} \times 0.20 \\ &= \$373\,333\end{aligned}$$

Calculation of the CCR capital charge

Question 2.b

The default probability of the counterparty is estimated at 1%. Calculate then the capital charge for counterparty credit risk of this OTC contract[a].

[a] We will take a value of 70% for the LGD parameter and a value of 20% for the default correlation. We can also use the approximations $-1.06 \approx -1$ and $\Phi(-1) \approx 16\%$.

Calculation of the CCR capital charge

While the bank uses the FIRB approach, the required capital is:

$$\mathcal{K} = \mathrm{EAD} \times \mathbb{E}\left[\mathrm{LGD}\right] \times \left(\Phi\left(\frac{\Phi^{-1}\left(\mathrm{PD}\right) + \sqrt{\rho}\Phi^{-1}\left(99.9\%\right)}{\sqrt{1-\rho}}\right) - \mathrm{PD}\right)$$

When ρ is equal to 20%, we have:

$$\begin{aligned}\frac{\Phi^{-1}\left(\mathrm{PD}\right) + \sqrt{\rho}\Phi^{-1}\left(99.9\%\right)}{\sqrt{1-\rho}} &= \frac{-2.33 + \sqrt{0.20} \times 3.09}{\sqrt{1-0.20}} \\ &= -1.06\end{aligned}$$

By using the approximations $-1.06 \simeq 1$ and $\Phi\left(-1\right) \simeq 0.16$, we obtain:

$$\begin{aligned}\mathcal{K} &= 373\,333 \times 0.70 \times \left(0.16 - 0.01\right) \\ &= \$39\,200\end{aligned}$$

The required capital of this OTC product for counterparty credit risk is then equal to \$39 200.

Calculation of CVA and DVA measures

We consider an OTC contract with maturity T between Bank A and Bank B. We denote by $\mathrm{MtM}(t)$ the risk-free mark-to-market of Bank A. The current date is set to $t = 0$ and we assume that:

$$\mathrm{MtM}(t) = N \cdot \sigma \cdot \sqrt{t} \cdot X$$

where N is the notional of the OTC contract, σ is the volatility of the underlying asset and X is a random variable, which is defined on the support $[-1, 1]$ and whose density function is:

$$f(x) = \frac{1}{2}$$

Calculation of CVA and DVA measures

Question 1

Define the concept of positive exposure $e^{+}(t)$. Show that the cumulative distribution function $\mathbf{F}_{[0,t]}$ of $e^{+}(t)$ has the following expression:

$$\mathbf{F}_{[0,t]}(x) = \mathbb{1}\left\{0 \leq x \leq \sigma\sqrt{t}\right\} \cdot \left(\frac{1}{2} + \frac{x}{2 \cdot N \cdot \sigma \cdot \sqrt{t}}\right)$$

where $\mathbf{F}_{[0,t]}(x) = 0$ if $x \leq 0$ and $\mathbf{F}_{[0,t]}(x) = 1$ if $x \geq \sigma\sqrt{t}$.

Calculation of CVA and DVA measures

The positive exposure $e^{+}(t)$ is the maximum between zero and the mark-to-market value:

$$\begin{aligned} e^{+}(t) &= \max\left(0, \mathrm{MtM}(t)\right) \\ &= \max\left(0, N\sigma\sqrt{t}X\right) \end{aligned}$$

We have:

$$\begin{aligned} \mathbf{F}_{[0,t]}(x) &= \Pr\left\{e^{+}(t) \leq x\right\} \\ &= \Pr\left\{\max\left(0, N\sigma\sqrt{t}X\right) \leq x\right\} \end{aligned}$$

We notice that:

$$\max\left(0, N\sigma\sqrt{t}X\right) = \begin{cases} 0 & \text{if } X \leq 0 \\ N\sigma\sqrt{t}X & \text{otherwise} \end{cases}$$

Calculation of CVA and DVA measures

By assuming that $x \in \left[0, N\sigma\sqrt{t}\right]$, we deduce that:

$$\begin{aligned}
\mathbf{F}_{[0,t]}(x) &= \Pr\left\{e^{+}(t) \leq x, X \leq 0\right\} + \Pr\left\{e^{+}(t) \leq x, X > 0\right\} \\
&= \Pr\{0 \leq x, X \leq 0\} + \Pr\left\{N\sigma\sqrt{t}X \leq x, X > 0\right\} \\
&= \frac{1}{2} + \frac{1}{2}\Pr\left\{N\sigma\sqrt{t}U \leq x\right\} \\
&= \frac{1}{2} + \frac{1}{2}\Pr\left\{U \leq \frac{x}{N\sigma\sqrt{t}}\right\}
\end{aligned}$$

where U is the standard uniform random variable. We finally obtain the following expression:

$$\mathbf{F}_{[0,t]}(x) = \frac{1}{2} + \frac{x}{2N\sigma\sqrt{t}}$$

If $x \leq 0$ or $x \geq N\sigma\sqrt{t}$, it is easy to show that $\mathbf{F}_{[0,t]}(x) = 0$ and $\mathbf{F}_{[0,t]}(x) = 1$.

Calculation of CVA and DVA measures

Question 2

Deduce the value of the expected positive exposure $\mathrm{EpE}(t)$.

Calculation of CVA and DVA measures

The expected positive exposure $\operatorname{EpE}(t)$ is defined as follows:

$$\operatorname{EpE}(t) = \mathbb{E}\left[e^{+}(t)\right]$$

Using the expression of $\mathbf{F}_{[0,t]}(x)$, it follows that the density function of $e^{+}(t)$ is equal to:

$$\begin{aligned} f_{[0,t]}(x) &= \frac{\partial\,\mathbf{F}_{[0,t]}(x)}{\partial\,x} \\ &= \frac{1}{2N\sigma\sqrt{t}} \end{aligned}$$

Calculation of CVA and DVA measures

We deduce that:

$$\begin{aligned}
\operatorname{EpE}(t) &= \int_0^{N\sigma\sqrt{t}} x f_{[0,t]}(x)\,\mathrm{d}x \\
&= \int_0^{N\sigma\sqrt{t}} \frac{x}{2N\sigma\sqrt{t}}\,\mathrm{d}x \\
&= \left[\frac{x^2}{4N\sigma\sqrt{t}}\right]_0^{N\sigma\sqrt{t}} \\
&= \frac{N\sigma\sqrt{t}}{4}
\end{aligned}$$

Calculation of CVA and DVA measures

Question 3

We note $\boldsymbol{\mathcal{R}}_B$ the fixed and constant recovery rate of Bank B. Give the mathematical expression of the CVA.

Calculation of CVA and DVA measures

By definition, we have:

$$\mathrm{CVA} = (1 - \boldsymbol{\mathcal{R}}_B) \times \int_0^T -B_0(t)\,\mathrm{EpE}(t)\;\mathrm{d}\mathbf{S}_B(t)$$

Calculation of CVA and DVA measures

Question 4

By using the definition of the lower incomplete gamma function $\gamma(s, x)$, show that the CVA is equal to:

$$\mathrm{CVA} = \frac{N \cdot (1 - \mathcal{R}_B) \cdot \sigma \cdot \gamma\left(\frac{3}{2}, \lambda_B T\right)}{4\sqrt{\lambda_B}}$$

when the default time of Bank B is exponential with parameter λ_B and interest rates are equal to zero.

Calculation of CVA and DVA measures

The interest rates are equal to zero meaning that $B_0(t) = 1$. Moreover, we have $\mathbf{S}_B(t) = e^{-\lambda_B t}$. We deduce that:

$$\begin{aligned}
\text{CVA} &= (1 - \mathcal{R}_B) \times \int_0^T \frac{N\sigma\sqrt{t}}{4} \lambda_B e^{-\lambda_B t} \, \mathrm{d}t \\
&= \frac{N\lambda_B (1 - \mathcal{R}_B) \sigma}{4} \int_0^T \sqrt{t} e^{-\lambda_B t} \, \mathrm{d}t
\end{aligned}$$

The definition of the incomplete gamma function is:

$$\gamma(s, x) = \int_0^x t^{s-1} e^{-t} \, \mathrm{d}t$$

Calculation of CVA and DVA measures

By considering the change of variable $y = \lambda_B t$, we obtain:

$$\begin{aligned}
\int_0^T \sqrt{t} e^{-\lambda_B t} \, \mathrm{d}t &= \int_0^{\lambda_B T} \sqrt{\frac{y}{\lambda_B}} e^{-y} \frac{\mathrm{d}y}{\lambda_B} \\
&= \frac{1}{\lambda_B^{3/2}} \int_0^{\lambda_B T} y^{3/2-1} e^{-y} \, \mathrm{d}y \\
&= \frac{\gamma\left(\frac{3}{2}, \lambda_B T\right)}{\lambda_B^{3/2}}
\end{aligned}$$

It follows that:

$$\mathrm{CVA} = \frac{N\left(1 - \mathcal{R}_B\right)\sigma\gamma\left(\frac{3}{2}, \lambda_B T\right)}{4\sqrt{\lambda_B}}$$

Calculation of CVA and DVA measures

Question 5

Comment on this result.

Calculation of CVA and DVA measures

The CVA is proportional to the notional N of the OTC contract, the loss given default $(1 - \mathcal{R}_B)$ of the counterparty and the volatility σ of the underlying asset. It is an increasing function of the maturity T because we have $\gamma\left(\frac{3}{2}, \lambda_B T_2\right) > \gamma\left(\frac{3}{2}, \lambda_B T_1\right)$ when $T_2 > T_1$. If the maturity is not very large (less than 10 years), the CVA is an increasing function of the default intensity λ_B.

Calculation of CVA and DVA measures

The limit cases are[33]:

$$\lim_{\lambda_B \to \infty} \mathrm{CVA} = \lim_{\lambda_B \to \infty} \frac{N\left(1-\mathcal{R}_B\right)\sigma\gamma\left(\frac{3}{2}, \lambda_B T\right)}{4\sqrt{\lambda_B}} = 0$$

and:

$$\lim_{T \to \infty} \mathrm{CVA} = \frac{N\left(1-\mathcal{R}_B\right)\sigma\Gamma\left(\frac{3}{2}\right)}{4\sqrt{\lambda_B}}$$

When the counterparty has a high default intensity, meaning that the default is imminent, the CVA is equal to zero because the mark-to-market value is close to zero. When the maturity is large, the CVA is a decreasing function of the intensity λ_B. Indeed, the probability to observe a large mark-to-market in the future increases when the default time is very far from the current date. We have illustrated these properties in Figure **??** with the following numerical values: $N = \$1$ mn, $\mathcal{R}_B = 40\%$ and $\sigma = 30\%$.

[33] We have $\lim_{x \to \infty} \gamma\left(s, x\right) = \Gamma\left(s\right)$.

Calculation of CVA and DVA measures

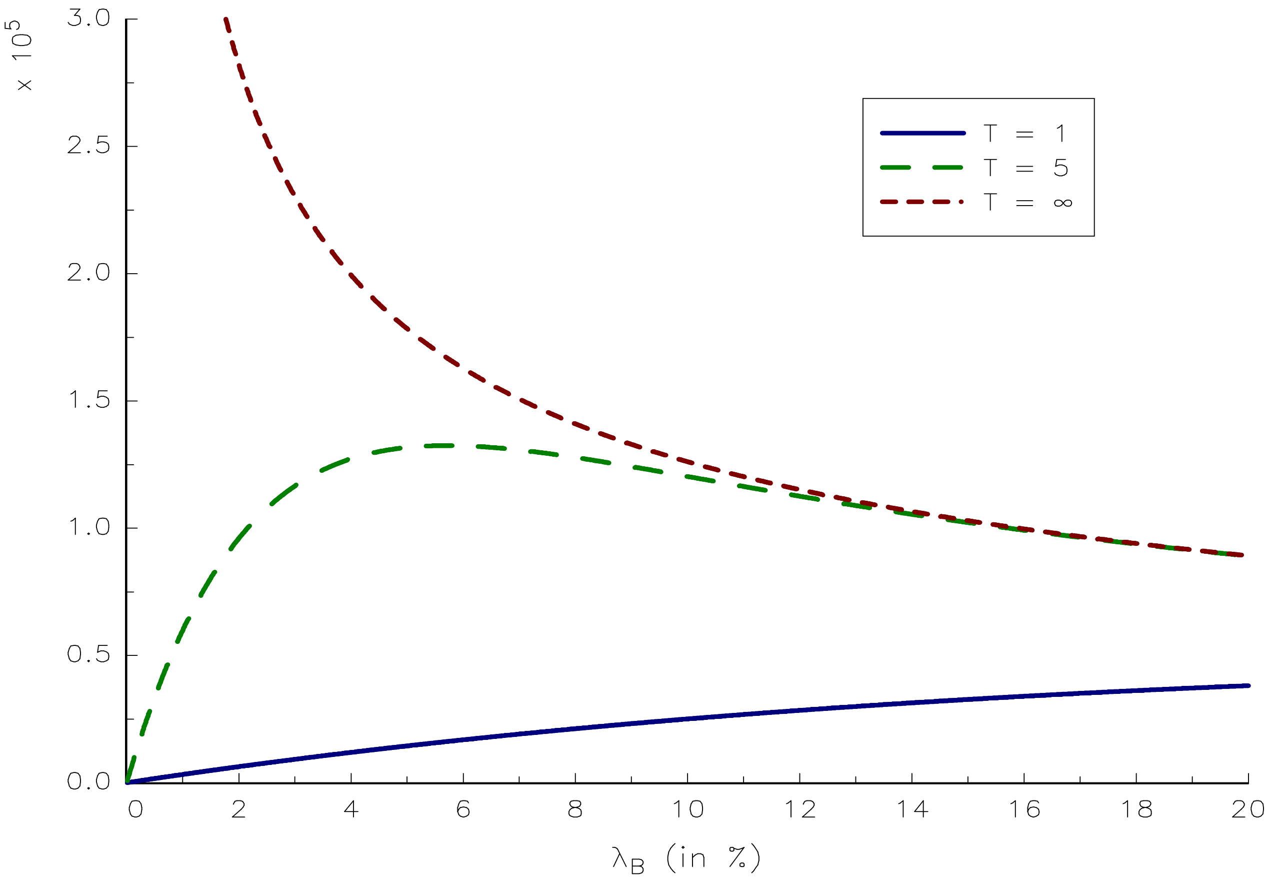


Figure: Evolution of the CVA with respect to maturity T and intensity λ_B

Calculation of CVA and DVA measures

Question 6

By assuming that the default time of Bank A is exponential with parameter λ_A, deduce the value of the DVA without additional computations.

Calculation of CVA and DVA measures

We notice that the mark-to-market is perfectly symmetric about 0. We deduce that the expected negative exposure $\mathrm{EnE}(t)$ is equal to the expected positive exposure $\mathrm{EpE}(t)$. It follows that the DVA is equal to:

$$\mathrm{DVA} = \frac{N(1 - \boldsymbol{\mathcal{R}}_A)\,\sigma\gamma\left(\frac{3}{2}, \lambda_A T\right)}{4\sqrt{\lambda_A}}$$

Course 2023-2024 in Financial Risk Management
Tutorial Session 4

Thierry Roncalli*

*Amundi Asset Management[34]

*University of Paris-Saclay

September 2023

[34] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Tutorial Session 1: Market Risk
- Tutorial Session 2: Credit Risk
- Tutorial Session 3: Counterparty Credit Risk and Collateral Risk
- **Tutorial Session 4: Operational Risk & Asset Liability Management Risk**
- Tutorial Session 5: Copulas, EVT & Stress Testing

Estimation of the loss severity distribution

Exercise

We consider a sample of n individual losses $\{x_1, \ldots, x_n\}$. We assume that they can be described by different probability distributions:

(i) X follows a log-normal distribution $\mathcal{LN}(\mu, \sigma^2)$.

(ii) X follows a Pareto distribution $\mathcal{P}(\alpha, x^-)$ defined by:

$$\Pr\{X \leq x\} = 1 - \left(\frac{x}{x_-}\right)^{-\alpha}$$

with $x \geq x_-$ and $\alpha > 0$.

(iii) X follows a gamma distribution $\Gamma(\alpha, \beta)$ defined by:

$$\Pr\{X \leq x\} = \int_0^x \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \, \mathrm{d}t$$

with $x \geq 0$, $\alpha > 0$ and $\beta > 0$.

(iv) The natural logarithm of the loss X follows a gamma distribution: $\ln X \sim \Gamma(\alpha; \beta)$.

Estimation of the loss severity distribution

Question 1

We consider the case *(i)*.

(i) X follows a log-normal distribution $\mathcal{LN}\left(\mu, \sigma^2\right)$.

Estimation of the loss severity distribution

Question 1.a

Show that the probability density function is:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right)$$

Estimation of the loss severity distribution

The density of the Gaussian distribution $Y \sim \mathcal{N}\left(\mu, \sigma^2\right)$ is:

$$g\left(y\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right)$$

Let $X \sim \mathcal{LN}\left(\mu, \sigma^2\right)$. We have $X = \exp Y$. It follows that:

$$f\left(x\right) = g\left(y\right)\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|$$

with $y = \ln x$. We deduce that:

$$\begin{aligned} f\left(x\right) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) \times \frac{1}{x} \\ &= \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right) \end{aligned}$$

Estimation of the loss severity distribution

Question 1.b

Calculate the two first moments of X. Deduce the orthogonal conditions of the generalized method of moments.

Estimation of the loss severity distribution

For $m \geq 1$, the non-centered moment is equal to:

$$\mathbb{E}\left[X^m\right] = \int_0^\infty x^m \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right) \mathrm{d}x$$

Estimation of the loss severity distribution

By considering the change of variables $y = \sigma^{-1}(\ln x - \mu)$ and $z = y - m\sigma$, we obtain:

$$\begin{aligned}
\mathbb{E}\left[X^m\right] &= \int_{-\infty}^{\infty} e^{m\mu + m\sigma y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\,\mathrm{d}y \\
&= e^{m\mu} \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2 + m\sigma y}\,\mathrm{d}y \\
&= e^{m\mu} \times e^{\frac{1}{2}m^2\sigma^2} \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y - m\sigma)^2}\,\mathrm{d}y \\
&= e^{m\mu + \frac{1}{2}m^2\sigma^2} \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right)\mathrm{d}z \\
&= e^{m\mu + \frac{1}{2}m^2\sigma^2}
\end{aligned}$$

Estimation of the loss severity distribution

We deduce that:

$$\mathbb{E}[X] = e^{\mu + \frac{1}{2}\sigma^2}$$

and:

$$\begin{array}{rcl} \mathrm{var}(X) & = & \mathbb{E}\left[X^2\right] - \mathbb{E}^2[X] \\ & = & e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} \\ & = & e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right) \end{array}$$

We can estimate the parameters μ and σ with the generalized method of moments by using the following empirical moments:

$$\left\{ \begin{array}{l} h_{i,1}(\mu,\sigma) = x_i - e^{\mu + \frac{1}{2}\sigma^2} \\ h_{i,2}(\mu,\sigma) = \left(x_i - e^{\mu+\frac{1}{2}\sigma^2}\right)^2 - e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right) \end{array} \right.$$

Estimation of the loss severity distribution

Question 1.c

Find the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$.

Estimation of the loss severity distribution

The log-likelihood function of the sample $\{x_1, \ldots, x_n\}$ is:

$$\begin{aligned}\ell(\mu, \sigma) &= \sum_{i=1}^{n} \ln f(x_i) \\ &= -\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln 2\pi - \sum_{i=1}^{n} \ln x_i - \frac{1}{2} \sum_{i=1}^{n} \left(\frac{\ln x_i - \mu}{\sigma} \right)^2\end{aligned}$$

To find the ML estimators $\hat{\mu}$ and $\hat{\sigma}$, we can proceed in two different way.

Estimation of the loss severity distribution

#1 $X \sim \mathcal{LN}\left(\mu, \sigma^2\right)$ implies that $Y = \ln X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. We know that the ML estimators $\hat{\mu}$ and $\hat{\sigma}$ associated to Y are:

$$\begin{aligned} \hat{\mu} &= \frac{1}{n}\sum_{i=1}^{n} y_i \\ \hat{\sigma} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(y_i - \hat{\mu}\right)^2} \end{aligned}$$

We deduce that the ML estimators $\hat{\mu}$ and $\hat{\sigma}$ associated to the sample $\{x_1, \ldots, x_n\}$ are:

$$\begin{aligned} \hat{\mu} &= \frac{1}{n}\sum_{i=1}^{n} \ln x_i \\ \hat{\sigma} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(\ln x_i - \hat{\mu}\right)^2} \end{aligned}$$

Estimation of the loss severity distribution

#2 We maximize the log-likelihood function. The first-order conditions are $\partial_\mu \boldsymbol{\ell}(\mu,\sigma) = 0$ and $\partial_\sigma \boldsymbol{\ell}(\mu,\sigma) = 0$. We deduce that:

$$\partial_\mu \boldsymbol{\ell}(\mu,\sigma) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_i - \mu) = 0$$

and:

$$\partial_\sigma \boldsymbol{\ell}(\mu,\sigma) = -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(\ln x_i - \mu)^2}{\sigma^3} = 0$$

We finally obtain:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i$$

and:

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2}$$

Estimation of the loss severity distribution

Question 2

We consider the case *(ii)*.

(ii) X follows a Pareto distribution $\mathcal{P}\left(\alpha, x^{-}\right)$ defined by:

$$\Pr\left\{X \leq x\right\} = 1 - \left(\frac{x}{x_{-}}\right)^{-\alpha}$$

with $x \geq x_{-}$ and $\alpha > 0$.

Estimation of the loss severity distribution

Question 2.a

Calculate the two first moments of X. Deduce the GMM conditions for estimating the parameter α.

Estimation of the loss severity distribution

The probability density function is:

$$
\begin{aligned}
f(x) &= \frac{\partial \Pr\{X \leq x\}}{\partial x} \\
&= \alpha \frac{x^{-(\alpha+1)}}{x_-^{-\alpha}}
\end{aligned}
$$

For $m \geq 1$, we have:

$$
\begin{aligned}
\mathbb{E}[X^m] &= \int_{x_-}^{\infty} x^m \alpha \frac{x^{-(\alpha+1)}}{x_-^{-\alpha}} \,\mathrm{d}x \\
&= \frac{\alpha}{x_-^{-\alpha}} \int_{x_-}^{\infty} x^{m-\alpha-1} \,\mathrm{d}x \\
&= \frac{\alpha}{x_-^{-\alpha}} \left[\frac{x^{m-\alpha}}{m-\alpha} \right]_{x_-}^{\infty} \\
&= \frac{\alpha}{\alpha - m} x_-^m
\end{aligned}
$$

Estimation of the loss severity distribution

We deduce that:

$$\mathbb{E}[X] = \frac{\alpha}{\alpha - 1} x_-$$

and:

$$\begin{aligned}
\operatorname{var}(X) &= \mathbb{E}\left[X^2\right] - \mathbb{E}^2[X] \\
&= \frac{\alpha}{\alpha - 2} x_-^2 - \left(\frac{\alpha}{\alpha - 1} x_-\right)^2 \\
&= \frac{\alpha}{(\alpha - 1)^2 (\alpha - 2)} x_-^2
\end{aligned}$$

Estimation of the loss severity distribution

We can then estimate the parameter α by considering the following empirical moments:

$$\begin{aligned} h_{i,1}(\alpha) &= x_i - \frac{\alpha}{\alpha - 1} x_- \\ h_{i,2}(\alpha) &= \left(x_i - \frac{\alpha}{\alpha - 1} x_- \right)^2 - \frac{\alpha}{(\alpha - 1)^2 (\alpha - 2)} x_-^2 \end{aligned}$$

The generalized method of moments can consider either the first moment $h_{i,1}(\alpha)$, the second moment $h_{i,2}(\alpha)$ or the joint moments $(h_{i,1}(\alpha), h_{i,2}(\alpha))$. In the first case, the estimator is:

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i - n x_-}$$

Estimation of the loss severity distribution

Question 2.b

Find the maximum likelihood estimator $\hat{\alpha}$.

Estimation of the loss severity distribution

The log-likelihood function is:

$$\ell(\alpha) = \sum_{i=1}^{n} \ln f(x_i) = n \ln \alpha - (\alpha + 1) \sum_{i=1}^{n} \ln x_i + n\alpha \ln x_-$$

The first-order condition is:

$$\partial_\alpha \ell(\alpha) = \frac{n}{\alpha} - \sum_{i=1}^{n} \ln x_i + \sum_{i=1}^{n} \ln x_- = 0$$

We deduce that:

$$n = \alpha \sum_{i=1}^{n} \ln \frac{x_i}{x_-}$$

The ML estimator is then:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} (\ln x_i - \ln x_-)}$$

Estimation of the loss severity distribution

Question 3

We consider the case *(iii)*. Write the log-likelihood function associated to the sample of individual losses $\{x_1, \ldots, x_n\}$. Deduce the first-order conditions of the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$.

(iii) X follows a gamma distribution $\Gamma(\alpha, \beta)$ defined by:

$$\Pr\{X \leq x\} = \int_0^x \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \, \mathrm{d}t$$

with $x \geq 0$, $\alpha > 0$ and $\beta > 0$.

Estimation of the loss severity distribution

The probability density function of *(iii)* is:

$$f(x) = \frac{\partial \Pr\{X \leq x\}}{\partial x} = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

It follows that the log-likelihood function is:

$$\ell(\alpha, \beta) = \sum_{i=1}^{n} \ln f(x_i) = -n \ln \Gamma(\alpha) + n\alpha \ln \beta + (\alpha - 1) \sum_{i=1}^{n} \ln x_i - \beta \sum_{i=1}^{n} x_i$$

The first-order conditions $\partial_{\alpha} \ell(\alpha, \beta) = 0$ and $\partial_{\beta} \ell(\alpha, \beta) = 0$ imply that:

$$n \left(\ln \beta - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \right) + \sum_{i=1}^{n} \ln x_i = 0$$

and:

$$n \frac{\alpha}{\beta} - \sum_{i=1}^{n} x_i = 0$$

Estimation of the loss severity distribution

Question 4

We consider the case *(iv)*. Show that the probability density function of X is:

$$f(x) = \frac{\beta^{\alpha} (\ln x)^{\alpha-1}}{\Gamma(\alpha) x^{\beta+1}}$$

What is the support of this probability density function? Write the log-likelihood function associated to the sample of individual losses $\{x_1, \ldots, x_n\}$.

(iv) The natural logarithm of the loss X follows a gamma distribution: $\ln X \sim \Gamma(\alpha; \beta)$.

Estimation of the loss severity distribution

Let $Y \sim \Gamma(\alpha, \beta)$ and $X = \exp Y$. We have:

$$f_X(x)\,|\mathrm{d}x| = f_Y(y)\,|\mathrm{d}y|$$

where f_X and f_Y are the probability density functions of X and Y. We deduce that:

$$\begin{aligned} f_X(x) &= \frac{\beta^\alpha y^{\alpha-1} e^{-\beta y}}{\Gamma(\alpha)} \times \frac{1}{e^y} \\ &= \frac{\beta^\alpha (\ln x)^{\alpha-1} e^{-\beta \ln x}}{x\Gamma(\alpha)} \\ &= \frac{\beta^\alpha (\ln x)^{\alpha-1}}{\Gamma(\alpha) x^{\beta+1}} \end{aligned}$$

The support of this probability density function is $[0, +\infty)$.

Estimation of the loss severity distribution

The log-likelihood function associated to the sample of individual losses $\{x_1, \ldots, x_n\}$ is:

$$\begin{aligned}\ell(\alpha, \beta) &= \sum_{i=1}^{n} \ln f(x_i) \\ &= -n \ln \Gamma(\alpha) + n\alpha \ln \beta + (\alpha - 1) \sum_{i=1}^{n} \ln(\ln x_i) - (\beta + 1) \sum_{i=1}^{n} \ln x_i\end{aligned}$$

Estimation of the loss severity distribution

Question 5

We now assume that the losses $\{x_1, \ldots, x_n\}$ have been collected beyond a threshold H meaning that $X \geq H$.

Estimation of the loss severity distribution

Question 5.a

What becomes the generalized method of moments in the case *(i)*.

(i) X follows a log-normal distribution $\mathcal{LN}\left(\mu, \sigma^2\right)$.

Estimation of the loss severity distribution

Using Bayes' formula, we have:

$$\begin{aligned}
\Pr\{X \leq x \mid X \geq H\} &= \frac{\Pr\{H \leq X \leq x\}}{\Pr\{X \geq H\}} \\
&= \frac{\mathbf{F}(x) - \mathbf{F}(H)}{1 - \mathbf{F}(H)}
\end{aligned}$$

where $\mathbf{F}$ is the cdf of X. We deduce that the conditional probability density function is:

$$\begin{aligned}
f(x \mid X \geq H) &= \partial_x \Pr\{X \leq x \mid X \geq H\} \\
&= \frac{f(x)}{1 - \mathbf{F}(H)} \times \mathbb{1}\{x \geq H\}
\end{aligned}$$

Estimation of the loss severity distribution

For the log-normal probability distribution, we obtain:

$$\begin{aligned} f(x \mid X \geq H) &= \frac{1}{1 - \Phi\left(\frac{\ln H - \mu}{\sigma}\right)} \times \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} \mathrm{d}x \\ &= \varphi \times \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2} \mathrm{d}x \end{aligned}$$

Estimation of the loss severity distribution

We note $\mathcal{M}_m(\mu,\sigma)$ the conditional moment $\mathbb{E}[X^m \mid X \geq H]$. We have:

$$\begin{aligned}
\mathcal{M}_m(\mu,\sigma) &= \varphi \times \int_H^\infty \frac{x^{m-1}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2} \,\mathrm{d}x \\
&= \varphi \times \int_{\ln H}^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + mx} \,\mathrm{d}x \\
&= \varphi \times e^{m\mu+\frac{1}{2}m^2\sigma^2} \times \int_{\ln H}^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{\left(x-\left(\mu+m\sigma^2\right)\right)^2}{\sigma^2}} \,\mathrm{d}x \\
&= \frac{1-\Phi\left(\frac{\ln H-\mu-m\sigma^2}{\sigma}\right)}{1-\Phi\left(\frac{\ln H-\mu}{\sigma}\right)} e^{m\mu+\frac{1}{2}m^2\sigma^2}
\end{aligned}$$

Estimation of the loss severity distribution

The first two moments of $X \mid X \geq H$ are then:

$$\mathcal{M}_1(\mu, \sigma) = \mathbb{E}[X \mid X \geq H] = \frac{1 - \Phi\left(\frac{\ln H - \mu - \sigma^2}{\sigma}\right)}{1 - \Phi\left(\frac{\ln H - \mu}{\sigma}\right)} e^{\mu + \frac{1}{2}\sigma^2}$$

and:

$$\mathcal{M}_2(\mu, \sigma) = \mathbb{E}\left[X^2 \mid X \geq H\right] = \frac{1 - \Phi\left(\frac{\ln H - \mu - 2\sigma^2}{\sigma}\right)}{1 - \Phi\left(\frac{\ln H - \mu}{\sigma}\right)} e^{2\mu + 2\sigma^2}$$

Estimation of the loss severity distribution

We can therefore estimate μ and σ by considering the following empirical moments:

$$\left\{ \begin{array}{l} h_{i,1}\left(\mu,\sigma\right) = x_i - \mathcal{M}_1\left(\mu,\sigma\right) \\ h_{i,2}\left(\mu,\sigma\right) = \left(x_i - \mathcal{M}_1\left(\mu,\sigma\right)\right)^2 - \left(\mathcal{M}_2\left(\mu,\sigma\right) - \mathcal{M}_1^2\left(\mu,\sigma\right)\right) \end{array} \right.$$

Estimation of the loss severity distribution

Question 5.b

Calculate the maximum likelihood estimator $\hat{\alpha}$ in the case *(ii)*.

(ii) X follows a Pareto distribution $\mathcal{P}(\alpha, x^-)$ defined by:

$$\Pr\{X \leq x\} = 1 - \left(\frac{x}{x_-}\right)^{-\alpha}$$

with $x \geq x_-$ and $\alpha > 0$.

Estimation of the loss severity distribution

We have:

$$\begin{aligned} f(x \mid X \geq H) &= \frac{f(x)}{1 - \mathbf{F}(H)} \times \mathbb{1}\{x \geq H\} \\ &= \left(\alpha \frac{x^{-(\alpha+1)}}{x_{-}^{-\alpha}}\right) \bigg/ \left(\frac{H^{-\alpha}}{x_{-}^{-\alpha}}\right) \\ &= \alpha \frac{x^{-(\alpha+1)}}{H^{-\alpha}} \end{aligned}$$

The conditional probability function is then a Pareto distribution with the same parameter α but with a new threshold $x_{-} = H$. We can then deduce that the ML estimator $\hat{\alpha}$ is:

$$\hat{\alpha} = \frac{n}{\left(\sum_{i=1}^{n} \ln x_i\right) - n \ln H}$$

Estimation of the loss severity distribution

Question 5.c

Write the log-likelihood function in the case *(iii)*.

(iii) X follows a gamma distribution $\Gamma(\alpha, \beta)$ defined by:

$$\Pr\{X \leq x\} = \int_0^x \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \,\mathrm{d}t$$

with $x \geq 0$, $\alpha > 0$ and $\beta > 0$.

Estimation of the loss severity distribution

The conditional probability density function is:

$$\begin{aligned}
f(x \mid X \geq H) &= \frac{f(x)}{1 - \mathbf{F}(H)} \times \mathbb{1}\{x \geq H\} \\
&= \left(\frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}\right) \Big/ \int_H^{\infty} \frac{\beta^{\alpha} t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \,\mathrm{d}t \\
&= \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\int_H^{\infty} \beta^{\alpha} t^{\alpha-1} e^{-\beta t} \,\mathrm{d}t}
\end{aligned}$$

We deduce that the log-likelihood function is:

$$\begin{aligned}
\ell(\alpha, \beta) &= n\alpha \ln \beta - n \ln \left(\int_H^{\infty} \beta^{\alpha} t^{\alpha-1} e^{-\beta t} \,\mathrm{d}t\right) + \\
&\quad (\alpha - 1) \sum_{i=1}^{n} \ln x_i - \beta \sum_{i=1}^{n} x_i
\end{aligned}$$

Estimation of the loss frequency distribution

Exercise

We consider a dataset of individual losses $\{x_1, \ldots, x_n\}$ corresponding to a sample of T annual loss numbers $\{N_{Y_1}, \ldots, N_{Y_T}\}$. This implies that:

$$\sum_{t=1}^{T} N_{Y_t} = n$$

If we measure the number of losses per quarter $\{N_{Q_1}, \ldots, N_{Q_{4T}}\}$, we use the notation:

$$\sum_{t=1}^{4T} N_{Q_t} = n$$

Estimation of the loss frequency distribution

Question 1

We assume that the annual number of losses follows a Poisson distribution $\mathcal{P}(\lambda_Y)$. Calculate the maximum likelihood estimator $\hat{\lambda}_Y$ associated to the sample $\{N_{Y_1}, \ldots, N_{Y_T}\}$.

Estimation of the loss frequency distribution

We have:

$$\Pr\{N = n\} = e^{-\lambda_Y} \frac{\lambda_Y^n}{n!}$$

We deduce that the expression of the log-likelihood function is:

$$\ell(\lambda_Y) = \sum_{t=1}^{T} \ln \Pr\{N = N_{Y_t}\} = -\lambda_Y T + \left(\sum_{t=1}^{T} N_{Y_t}\right) \ln \lambda_Y - \sum_{t=1}^{T} \ln(N_{Y_t}!)$$

The first-order condition is:

$$\frac{\partial\, \ell(\lambda_Y)}{\partial\, \lambda_Y} = -T + \frac{1}{\lambda_Y}\left(\sum_{t=1}^{T} N_{Y_t}\right) = 0$$

We deduce that the ML estimator is:

$$\hat{\lambda}_Y = \frac{1}{T}\sum_{t=1}^{T} N_{Y_t} = \frac{n}{T}$$

Estimation of the loss frequency distribution

Question 2

We assume that the quarterly number of losses follows a Poisson distribution $\mathcal{P}(\lambda_Q)$. Calculate the maximum likelihood estimator $\hat{\lambda}_Q$ associated to the sample $\{N_{Q_1}, \ldots, N_{Q_{4T}}\}$.

Estimation of the loss frequency distribution

Using the same arguments, we obtain:

$$\hat{\lambda}_Q = \frac{1}{4T}\sum_{t=1}^{4T} N_{Q_t} = \frac{n}{4T} = \frac{\hat{\lambda}_Y}{4}$$

Estimation of the loss frequency distribution

Question 3

What is the impact of considering a quarterly or annual basis on the computation of the capital charge?

Estimation of the loss frequency distribution

Considering a quarterly or annual basis has no impact on the capital charge. Indeed, the capital charge is computed with a one-year time horizon. If we use a quarterly basis, we have to find the distribution of the annual loss number. In this case, the annual loss number is the sum of the four quarterly loss numbers:

$$N_Y = N_{Q_1} + N_{Q_2} + N_{Q_3} + N_{Q_4}$$

We know that each quarterly loss number follows a Poisson distribution $\mathcal{P}\left(\hat{\lambda}_Q\right)$ and that they are independent. Because the Poisson distribution is infinitely divisible, we obtain:

$$N_{Q_1} + N_{Q_2} + N_{Q_3} + N_{Q_4} \sim \mathcal{P}\left(4\hat{\lambda}_Q\right)$$

We deduce that the annual loss number follows a Poisson distribution $\mathcal{P}\left(\hat{\lambda}_Y\right)$ in both cases.

Estimation of the loss frequency distribution

Question 4

What does this result become if we consider a method of moments based on the first moment?

Estimation of the loss frequency distribution

Since we have $\mathbb{E}\left[\mathcal{P}\left(\lambda\right)\right] = \lambda$, the MM estimator in the case of annual loss numbers is:

$$\hat{\lambda}_Y = \frac{1}{T}\sum_{t=1}^{T} N_{Y_t} = \frac{n}{T}$$

The MM estimator is exactly the ML estimator.

Estimation of the loss frequency distribution

Question 5

Same question if we consider a method of moments based on the second moment.

Estimation of the loss frequency distribution

Since we have $\operatorname{var}(\mathcal{P}(\lambda)) = \lambda$, the MM estimator in the case of annual loss numbers is:

$$\hat{\lambda}_Y = \frac{1}{T}\sum_{t=1}^{T} N_{Y_t}^2 - \frac{n^2}{T^2}$$

If we use a quarterly basis, we obtain:

$$\begin{aligned}\hat{\lambda}_Q &= \frac{1}{4}\left(\frac{1}{T}\sum_{t=1}^{4T} N_{Q_t}^2 - \frac{n^2}{4T^2}\right) \\ &\neq \frac{\hat{\lambda}_Y}{4}\end{aligned}$$

There is no reason that $\hat{\lambda}_Y = 4\hat{\lambda}_Q$ meaning that the capital charge will not be the same.

Computation of the amortization functions

Exercise

In what follows, we consider a debt instrument, whose remaining maturity is equal to m. We note t the current date and $T = t + m$ the maturity date.

Computation of the amortization functions

Question 1

We consider a bullet repayment debt. Define its amortization function $\mathbf{S}(t, u)$. Calculate the survival function $\mathbf{S}^{\star}(t, u)$ of the stock. Show that:

$$\mathbf{S}^{\star}(t, u) = \mathbb{1}\{t \leq u < t + m\} \cdot \left(1 - \frac{u - t}{m}\right)$$

in the case where the new production is constant. Comment on this result.

Computation of the amortization functions

By definition, we have:

$$\mathbf{S}(t,u) = \mathbb{1}\{t \leq u < t+m\} = \begin{cases} 1 & \text{if } u \in [t, t+m[\\ 0 & \text{otherwise} \end{cases}$$

This means that the survival function is equal to one when u is between the current date t and the maturity date $T = t + m$. When u reaches T, the outstanding amount is repaid, implying that $\mathbf{S}(t,T)$ is equal to zero. It follows that:

$$\begin{aligned} \mathbf{S}^{\star}(t,u) &= \frac{\int_{-\infty}^{t} \mathrm{NP}(s)\, \mathbf{S}(s,u)\, \mathrm{d}s}{\int_{-\infty}^{t} \mathrm{NP}(s)\, \mathbf{S}(s,t)\, \mathrm{d}s} \\ &= \frac{\int_{-\infty}^{t} \mathrm{NP}(s) \cdot \mathbb{1}\{s \leq u < s+m\}\, \mathrm{d}s}{\int_{-\infty}^{t} \mathrm{NP}(s) \cdot \mathbb{1}\{s \leq t < s+m\}\, \mathrm{d}s} \end{aligned}$$

Computation of the amortization functions

For the numerator, we have:

$$\begin{aligned} \mathbb{1}\{s \leq u < s + m\} = 1 \quad &\Rightarrow \quad u < s + m \\ &\Leftrightarrow \quad s > u - m \end{aligned}$$

and:

$$\int_{-\infty}^{t} \mathrm{NP}(s) \cdot \mathbb{1}\{s \leq u < s + m\}\, \mathrm{d}s = \int_{u-m}^{t} \mathrm{NP}(s)\, \mathrm{d}s$$

Computation of the amortization functions

For the denominator, we have:

$$\begin{aligned}\mathbb{1}\{s \leq t < s+m\} = 1 &\Rightarrow t < s+m \\ &\Leftrightarrow s > t-m\end{aligned}$$

and:

$$\int_{-\infty}^{t} \mathrm{NP}(s) \cdot \mathbb{1}\{s \leq t < s+m\} \,\mathrm{d}s = \int_{t-m}^{t} \mathrm{NP}(s) \,\mathrm{d}s$$

We deduce that:

$$\mathbf{S}^{\star}(t,u) = \mathbb{1}\{t \leq u < t+m\} \cdot \frac{\int_{u-m}^{t} \mathrm{NP}(s) \,\mathrm{d}s}{\int_{t-m}^{t} \mathrm{NP}(s) \,\mathrm{d}s}$$

Computation of the amortization functions

In the case where the new production is a constant, we have $\mathrm{NP}(s) = c$ and:

$$\begin{aligned}
\mathbf{S}^{\star}(t, u) &= \mathbb{1}\{t \leq u < t + m\} \cdot \frac{\int_{u-m}^{t} \mathrm{d}s}{\int_{t-m}^{t} \mathrm{d}s} \\
&= \mathbb{1}\{t \leq u < t + m\} \cdot \frac{\left[s\right]_{u-m}^{t}}{\left[s\right]_{t-m}^{t}} \\
&= \mathbb{1}\{t \leq u < t + m\} \cdot \left(\frac{t - u + m}{t - t + m}\right) \\
&= \mathbb{1}\{t \leq u < t + m\} \cdot \left(1 - \frac{u - t}{m}\right)
\end{aligned}$$

The survival function $\mathbf{S}^{\star}(t, u)$ corresponds to the case of a linear amortization.

Computation of the amortization functions

Question 2

Same question if we consider a debt instrument, whose amortization rate is constant.

Computation of the amortization functions

If the amortization is linear, we have:

$$\mathbf{S}(t,u) = \mathbb{1}\{t \leq u < t+m\} \cdot \left(1 - \frac{u-t}{m}\right)$$

We deduce that:

$$\mathbf{S}^{\star}(t,u) = \mathbb{1}\{t \leq u < t+m\} \cdot \frac{\int_{u-m}^{t} \mathrm{NP}(s)\left(1 - \frac{u-s}{m}\right)\,\mathrm{d}s}{\int_{t-m}^{t} \mathrm{NP}(s)\left(1 - \frac{t-s}{m}\right)\,\mathrm{d}s}$$

In the case where the new production is a constant, we obtain:

$$\mathbf{S}^{\star}(t,u) = \mathbb{1}\{t \leq u < t+m\} \cdot \frac{\int_{u-m}^{t} \left(1 - \frac{u-s}{m}\right)\,\mathrm{d}s}{\int_{t-m}^{t} \left(1 - \frac{t-s}{m}\right)\,\mathrm{d}s}$$

Computation of the amortization functions

For the numerator, we have:

$$\begin{aligned}
\int_{u-m}^{t}\left(1-\frac{u-s}{m}\right)\mathrm{d}s &= \left[s-\frac{su}{m}+\frac{s^2}{2m}\right]_{u-m}^{t} \\
&= \left(t-\frac{tu}{m}+\frac{t^2}{2m}\right)- \\
&\quad \left(u-m-\frac{u^2-mu}{m}+\frac{(u-m)^2}{2m}\right) \\
&= \left(t-\frac{tu}{m}+\frac{t^2}{2m}\right)-\left(u-\frac{m}{2}-\frac{u^2}{2m}\right) \\
&= \frac{m^2+u^2+t^2+2mt-2mu-2tu}{2m} \\
&= \frac{(m-u+t)^2}{2m}
\end{aligned}$$

Computation of the amortization functions

For the denominator, we use the previous result and we set $u = t$:

$$\begin{aligned}\int_{t-m}^{t} \left(1 - \frac{t-s}{m}\right) \mathrm{d}s &= \frac{(m - t + t)^2}{2m} \\ &= \frac{m}{2}\end{aligned}$$

Computation of the amortization functions

We deduce that:

$$\begin{aligned}
\mathbf{S}^{\star}(t,u) &= \mathbb{1}\left\{t \leq u < t+m\right\} \cdot \frac{\dfrac{(m-u+t)^2}{2m}}{\dfrac{m}{2}} \\
&= \mathbb{1}\left\{t \leq u < t+m\right\} \cdot \frac{(m-u+t)^2}{m^2} \\
&= \mathbb{1}\left\{t \leq u < t+m\right\} \cdot \left(1 - \frac{u-t}{m}\right)^2
\end{aligned}$$

The survival function $\mathbf{S}^{\star}(t,u)$ corresponds to the case of a parabolic amortization.

Computation of the amortization functions

Question 3

Same question if we assume[a] that the amortization function is exponential with parameter λ.

[a] By definition of the exponential amortization, we have $m = +\infty$.

Computation of the amortization functions

If the amortization is exponential, we have:

$$\mathbf{S}(t,u) = e^{-\int_t^u \lambda \,\mathrm{d}s} = e^{-\lambda(u-t)}$$

It follows that:

$$\mathbf{S}^{\star}(t,u) = \frac{\int_{-\infty}^{t} \mathrm{NP}(s)\, e^{-\lambda(u-s)}\,\mathrm{d}s}{\int_{-\infty}^{t} \mathrm{NP}(s)\, e^{-\lambda(t-s)}\,\mathrm{d}s}$$

In the case where the new production is a constant, we obtain:

$$\begin{aligned}
\mathbf{S}^{\star}(t,u) &= \frac{\int_{-\infty}^{t} e^{-\lambda(u-s)}\,\mathrm{d}s}{\int_{-\infty}^{t} e^{-\lambda(t-s)}\,\mathrm{d}s} \\
&= \frac{\left[\lambda^{-1}e^{-\lambda(u-s)}\right]_{-\infty}^{t}}{\left[\lambda^{-1}e^{-\lambda(t-s)}\right]_{-\infty}^{t}} \\
&= e^{-\lambda(u-t)} \\
&= \mathbf{S}(t,u)
\end{aligned}$$

The stock amortization function is equal to the flow amortization function.

Computation of the amortization functions

Question 4

Find the expression of $\mathcal{D}^{\star}(t)$ when the new production is constant.

Computation of the amortization functions

We recall that the liquidity duration is equal to:

$$\mathcal{D}(t) = \int_t^\infty (u - t) f(t, u) \, \mathrm{d}u$$

where $f(t, u)$ is the density function associated to the survival function $\mathbf{S}(t, u)$. For the stock, we have:

$$\mathcal{D}^\star(t) = \int_t^\infty (u - t) f^\star(t, u) \, \mathrm{d}u$$

where $f^\star(t, u)$ is the density function associated to the survival function $\mathbf{S}^\star(t, u)$:

$$f^\star(t, u) = \frac{\int_{-\infty}^t \mathrm{NP}(s) f(s, u) \, \mathrm{d}s}{\int_{-\infty}^t \mathrm{NP}(s) \mathbf{S}(s, t) \, \mathrm{d}s}$$

Computation of the amortization functions

In the case where the new production is constant, we obtain:

$$\mathcal{D}^{\star}(t) = \frac{\int_t^{\infty} (u-t) \int_{-\infty}^{t} f(s,u) \,\mathrm{d}s \,\mathrm{d}u}{\int_{-\infty}^{t} \mathbf{S}(s,t) \,\mathrm{d}s}$$

Since we have $\int_{-\infty}^{t} f(s,u) \,\mathrm{d}s = \mathbf{S}(t,u)$, we deduce that:

$$\mathcal{D}^{\star}(t) = \frac{\int_t^{\infty} (u-t) \,\mathbf{S}(t,u) \,\mathrm{d}u}{\int_{-\infty}^{t} \mathbf{S}(s,t) \,\mathrm{d}s}$$

Computation of the amortization functions

Question 5

Calculate the durations $\mathcal{D}(t)$ and $\mathcal{D}^{\star}(t)$ for the three previous cases.

Computation of the amortization functions

In the case of the bullet repayment debt, we have:

$$\mathcal{D}(t) = m$$

and:

$$\begin{aligned}
\mathcal{D}^{\star}(t) &= \frac{\int_t^{t+m} (u-t)\,\mathrm{d}u}{\int_{t-m}^{t} \mathrm{d}s} \\
&= \frac{\left[\frac{1}{2}(u-t)^2\right]_t^{t+m}}{\left[s\right]_{t-m}^{t}} \\
&= \frac{m}{2}
\end{aligned}$$

Computation of the amortization functions

In the case of the linear amortization, we have:

$$f(t, u) = \mathbb{1}\{t \leq u < t + m\} \cdot \frac{1}{m}$$

and:

$$\begin{aligned} \mathcal{D}(t) &= \int_t^{t+m} \frac{(u - t)}{m} \,\mathrm{d}u \\ &= \frac{1}{m} \left[\frac{1}{2} (u - t)^2 \right]_t^{t+m} \\ &= \frac{m}{2} \end{aligned}$$

Computation of the amortization functions

For the stock duration, we deduce that

$$
\begin{aligned}
\mathcal{D}^{\star}(t) &= \frac{\displaystyle\int_{t}^{t+m} (u-t)\left(1-\frac{u-t}{m}\right)\mathrm{d}u}{\displaystyle\int_{t-m}^{t}\left(1-\frac{t-s}{m}\right)\mathrm{d}s} \\
&= \frac{\displaystyle\int_{t}^{t+m}\left(u-t-\frac{u^2}{m}+2\frac{tu}{m}-\frac{t^2}{m}\right)\mathrm{d}u}{\displaystyle\int_{t-m}^{t}\left(1-\frac{t}{m}+\frac{s}{m}\right)\mathrm{d}s} \\
&= \frac{\left[\dfrac{u^2}{2}-tu-\dfrac{u^3}{3m}+\dfrac{tu^2}{m}-\dfrac{t^2u}{m}\right]_{t}^{t+m}}{\left[s-\dfrac{st}{m}+\dfrac{s^2}{2m}\right]_{t-m}^{t}}
\end{aligned}
$$

Computation of the amortization functions

The numerator is equal to:

$$\begin{aligned}
(*) &= \left[\frac{u^2}{2} - tu - \frac{u^3}{3m} + \frac{tu^2}{m} - \frac{t^2 u}{m}\right]_t^{t+m} \\
&= \frac{1}{6m}\left[3mu^2 - 6mtu - 2u^3 + 6tu^2 - 6t^2 u\right]_t^{t+m} \\
&= \frac{1}{6m}\left(m^3 - 3mt^2 - 2t^3\right) + \frac{1}{6m}\left(3mt^2 + 2t^3\right) \\
&= \frac{m^2}{6}
\end{aligned}$$

Computation of the amortization functions

The denominator is equal to:

$$
\begin{aligned}
(*) &= \left[s - \frac{st}{m} + \frac{s^2}{2m}\right]_{t-m}^{t} \\
&= \frac{1}{2m}\left[s^2 - 2s\left(t-m\right)\right]_{t-m}^{t} \\
&= \frac{1}{2m}\left(t^2 - 2t\left(t-m\right) - \left(t-m\right)^2 + 2\left(t-m\right)^2\right) \\
&= \frac{1}{2m}\left(t^2 - 2t^2 + 2mt + t^2 - 2mt + m^2\right) \\
&= \frac{m}{2}
\end{aligned}
$$

Computation of the amortization functions

We deduce that:

$$\mathcal{D}^{\star}(t) = \frac{m}{3}$$

Computation of the amortization functions

For the exponential amortization, we have:

$$f(t,u) = \lambda e^{-\lambda(u-t)}$$

and[35]:

$$\mathcal{D}(t) = \int_t^{\infty} (u-t)\,\lambda e^{-\lambda(u-t)}\,\mathrm{d}u = \int_0^{\infty} v\lambda e^{-\lambda v}\,\mathrm{d}v = \frac{1}{\lambda}$$

For the stock duration, we deduce that:

$$\mathcal{D}^{\star}(t) = \frac{\int_t^{\infty}(u-t)\,e^{-\lambda(u-t)}\,\mathrm{d}u}{\int_{-\infty}^{t} e^{-\lambda(t-s)}\,\mathrm{d}s} = \frac{\int_0^{\infty} v e^{-\lambda v}\,\mathrm{d}v}{\int_0^{\infty} e^{-\lambda v}\,\mathrm{d}v} = \frac{1}{\lambda}$$

We verify that $\mathcal{D}(t) = \mathcal{D}^{\star}(t)$ since we have demonstrated that $\mathbf{S}^{\star}(t,u) = \mathbf{S}(t,u)$.

[35] We use the change of variable $v = u - t$.

Computation of the amortization functions

Question 6

Calculate the corresponding dynamics $\mathrm{d}N(t)$.

Computation of the amortization functions

In the case of the bullet repayment debt, we have:

$$\mathrm{d}N(t) = (\mathrm{NP}(t) - \mathrm{NP}(t - m))\,\mathrm{d}t$$

Computation of the amortization functions

In the case of the linear amortization, we have:

$$f(s,t) = \frac{\mathbf{1}\{s \leq t < s+m\}}{m}$$

It follows that:

$$\begin{aligned}\int_{-\infty}^{t} \mathrm{NP}(s) f(s,t)\,\mathrm{d}s &= \frac{1}{m}\int_{-\infty}^{t} \mathbf{1}\{s \leq t < s+m\} \cdot \mathrm{NP}(s)\,\mathrm{d}s \\ &= \frac{1}{m}\int_{t-m}^{t} \mathrm{NP}(s)\,\mathrm{d}s\end{aligned}$$

We deduce that:

$$\mathrm{d}N(t) = \left(\mathrm{NP}(t) - \frac{1}{m}\int_{t-m}^{t} \mathrm{NP}(s)\,\mathrm{d}s\right)\mathrm{d}t$$

Computation of the amortization functions

For the exponential amortization, we have:

$$f(s,t) = \lambda e^{-\lambda(t-s)}$$

and:

$$\begin{aligned}\int_{-\infty}^{t} \mathrm{NP}(s) f(s,t)\,\mathrm{d}s &= \int_{-\infty}^{t} \mathrm{NP}(s)\,\lambda e^{-\lambda(t-s)}\,\mathrm{d}s \\ &= \lambda \int_{-\infty}^{t} \mathrm{NP}(s)\,e^{-\lambda(t-s)}\,\mathrm{d}s \\ &= \lambda N(t)\end{aligned}$$

We deduce that:

$$\mathrm{d}N(t) = (\mathrm{NP}(t) - \lambda N(t))\,\mathrm{d}t$$

Impact of prepayment

Exercise

We recall that the outstanding balance of a CAM (constant amortization mortgage) at time t is given by:

$$N(t) = \mathbf{1}\{t < m\} \cdot N_0 \cdot \frac{1 - e^{-i(m-t)}}{1 - e^{-im}}$$

where N_0 is the notional, i is the interest rate and m is the maturity.

Impact of prepayment

Question 1

Find the dynamics $\mathrm{d}N(t)$.

Impact of prepayment

We deduce that the dynamics of $N(t)$ is equal to:

$$\begin{aligned}
\mathrm{d}N(t) &= \mathbb{1}\{t < m\} \cdot N_0 \frac{-ie^{-i(m-t)}}{1 - e^{-im}} \,\mathrm{d}t \\
&= -ie^{-i(m-t)} \left(\mathbb{1}\{t < m\} \cdot N_0 \frac{1}{1 - e^{-im}} \right) \mathrm{d}t \\
&= -\frac{ie^{-i(m-t)}}{1 - e^{-i(m-t)}} N(t) \,\mathrm{d}t
\end{aligned}$$

Impact of prepayment

Question 2

We note $\tilde{N}(t)$ the modified outstanding balance that takes into account the prepayment risk. Let $\lambda_p(t)$ be the prepayment rate at time t. Write the dynamics of $\tilde{N}(t)$.

Impact of prepayment

The prepayment rate has a negative impact on $\mathrm{d}N(t)$ because it reduces the outstanding amount $N(t)$:

$$\mathrm{d}\tilde{N}(t) = -\frac{ie^{-i(m-t)}}{1-e^{-i(m-t)}}\tilde{N}(t)\,\mathrm{d}t - \lambda_p(t)\,\tilde{N}(t)\,\mathrm{d}t$$

Impact of prepayment

Question 3

Show that $\tilde{N}(t) = N(t)\,\mathbf{S}_p(t)$ where $\mathbf{S}_p(t)$ is the prepayment-based survival function.

Impact of prepayment

It follows that:

$$\mathrm{d}\ln\tilde{N}(t) = -\left(\frac{ie^{-i(m-t)}}{1-e^{-i(m-t)}} + \lambda_p(t)\right)\mathrm{d}t$$

and:

$$\begin{aligned}\ln\tilde{N}(t) - \ln\tilde{N}(0) &= \int_0^t \frac{-ie^{-i(m-s)}}{1-e^{-i(m-s)}}\,\mathrm{d}s - \int_0^t \lambda_p(s)\,\mathrm{d}s \\ &= \left[\ln\left(1-e^{-i(m-s)}\right)\right]_0^t - \int_0^t \lambda_p(s)\,\mathrm{d}s \\ &= \ln\left(\frac{1-e^{-i(m-t)}}{1-e^{-im}}\right) - \int_0^t \lambda_p(s)\,\mathrm{d}s\end{aligned}$$

Impact of prepayment

We deduce that:

$$\begin{aligned}\tilde{N}(t) &= \left(N_0 \frac{1-e^{-i(m-t)}}{1-e^{-im}}\right) e^{-\int_0^t \lambda_p(s)\,\mathrm{d}s} \\ &= N(t)\,\mathbf{S}_p(t)\end{aligned}$$

where $\mathbf{S}_p(t)$ is the survival function associated to the hazard rate $\lambda_p(t)$.

Impact of prepayment

Question 4

Calculate the liquidity duration $\tilde{\mathcal{D}}(t)$ associated to the outstanding balance $\tilde{N}(t)$ when the hazard rate of prepayments is constant and equal to λ_p.

Impact of prepayment

We have:

$$\tilde{N}(t,u) = \mathbf{1}\{t \leq u < t+m\} \cdot N(t) \frac{1-e^{-i(t+m-u)}}{1-e^{-im}} e^{-\lambda_p(u-t)}$$

this implies that:

$$\tilde{\mathbf{S}}(t,u) = \mathbf{1}\{t \leq u < t+m\} \cdot \frac{e^{-\lambda_p(u-t)} - e^{-im+(i-\lambda_p)(u-t)}}{1-e^{-im}}$$

and:

$$\tilde{f}(t,u) = \mathbf{1}\{t \leq u < t+m\} \cdot \frac{\lambda_p e^{-\lambda_p(u-t)} + (i-\lambda_p)\, e^{-im+(i-\lambda_p)(u-t)}}{1-e^{-im}}$$

Impact of prepayment

It follows that:

$$\begin{aligned}
\tilde{\mathcal{D}}(t) &= \frac{\lambda_p}{1-e^{-im}} \int_t^{t+m} (u-t)\, e^{-\lambda_p (u-t)}\,\mathrm{d}u + \\
&\quad \frac{(i-\lambda_p)\, e^{-im}}{1-e^{-im}} \int_t^{t+m} (u-t)\, e^{(i-\lambda_p)(u-t)}\,\mathrm{d}u \\
&= \frac{\lambda_p}{1-e^{-im}} \int_0^{m} v e^{-\lambda_p v}\,\mathrm{d}v + \frac{(i-\lambda_p)\, e^{-im}}{1-e^{-im}} \int_0^{m} v e^{(i-\lambda_p) v}\,\mathrm{d}v \\
&= \frac{\lambda_p}{1-e^{-im}} \left(\frac{m e^{-\lambda_p m}}{-\lambda_p} - \frac{e^{-\lambda_p m}-1}{\lambda_p^2} \right) + \\
&\quad \frac{(i-\lambda_p)\, e^{-im}}{1-e^{-im}} \left(\frac{m e^{(i-\lambda_p) m}}{(i-\lambda_p)} - \frac{e^{(i-\lambda_p) m}-1}{(i-\lambda_p)^2} \right) \\
&= \frac{1}{1-e^{-im}} \left(\frac{e^{-im}-e^{-\lambda_p m}}{i-\lambda_p} + \frac{1-e^{-\lambda_p m}}{\lambda_p} \right)
\end{aligned}$$

Impact of prepayment

because we have:

$$\begin{aligned}
\int_0^m v e^{\alpha v} \, \mathrm{d}v &= \left[\frac{v e^{\alpha v}}{\alpha}\right]_0^m - \int_0^m \frac{e^{\alpha v}}{\alpha} \, \mathrm{d}v \\
&= \left[\frac{v e^{\alpha v}}{\alpha}\right]_0^m - \left[\frac{e^{\alpha v}}{\alpha^2}\right]_0^m \\
&= \frac{m e^{\alpha m}}{\alpha} - \frac{e^{\alpha m} - 1}{\alpha^2}
\end{aligned}$$

Course 2023-2024 in Financial Risk Management
Tutorial Session 5

Thierry Roncalli*

*Amundi Asset Management[36]

*University of Paris-Saclay

September 2023

[36] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

Agenda

- Tutorial Session 1: Market Risk
- Tutorial Session 2: Credit Risk
- Tutorial Session 3: Counterparty Credit Risk and Collateral Risk
- Tutorial Session 4: Operational Risk & Asset Liability Management Risk
- **Tutorial Session 5: Copulas, EVT & Stress Testing**

The bivariate Pareto copula

Exercise

We consider the bivariate Pareto distribution:

$$\begin{aligned}\mathbf{F}(x_1, x_2) &= 1 - \left(\frac{\theta_1 + x_1}{\theta_1}\right)^{-\alpha} - \left(\frac{\theta_2 + x_2}{\theta_2}\right)^{-\alpha} + \\ &\quad \left(\frac{\theta_1 + x_1}{\theta_1} + \frac{\theta_2 + x_2}{\theta_2} - 1\right)^{-\alpha}\end{aligned}$$

where $x_1 \geq 0$, $x_2 \geq 0$, $\theta_1 > 0$, $\theta_2 > 0$ and $\alpha > 0$.

The bivariate Pareto copula

Question 1

Show that the marginal functions of $\mathbf{F}(x_1, x_2)$ correspond to univariate Pareto distributions.

The bivariate Pareto copula

We have:

$$\begin{aligned}
\mathbf{F}_1(x_1) &= \Pr\{X_1 \leq x_1\} \\
&= \Pr\{X_1 \leq x_1, X_2 \leq \infty\} \\
&= \mathbf{F}(x_1, \infty)
\end{aligned}$$

We deduce that:

$$\begin{aligned}
\mathbf{F}_1(x_1) &= 1 - \left(\frac{\theta_1 + x_1}{\theta_1}\right)^{-\alpha} - \left(\frac{\theta_2 + \infty}{\theta_2}\right)^{-\alpha} + \\
&\quad \left(\frac{\theta_1 + x_1}{\theta_1} + \frac{\theta_2 + \infty}{\theta_2} - 1\right)^{-\alpha} \\
&= 1 - \left(\frac{\theta_1 + x_1}{\theta_1}\right)^{-\alpha}
\end{aligned}$$

We conclude that $\mathbf{F}_1$ (and $\mathbf{F}_2$) is a Pareto distribution.

The bivariate Pareto copula

Question 2

Find the copula function associated to the bivariate Pareto distribution.

The bivariate Pareto copula

We have:

$$\mathbf{C}(u_1, u_2) = \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)$$

It follows that:

$$\begin{aligned} 1 - \left(\frac{\theta_1 + x_1}{\theta_1}\right)^{-\alpha} &= u_1 \\ \Leftrightarrow \quad \left(\frac{\theta_1 + x_1}{\theta_1}\right)^{-\alpha} &= 1 - u_1 \\ \Leftrightarrow \quad \frac{\theta_1 + x_1}{\theta_1} &= (1 - u_1)^{-1/\alpha} \end{aligned}$$

We deduce that:

$$\begin{aligned} \mathbf{C}(u_1, u_2) &= 1 - (1 - u_1) - (1 - u_2) + \\ & \quad \left((1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1\right)^{-\alpha} \\ &= u_1 + u_2 - 1 + \left((1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1\right)^{-\alpha} \end{aligned}$$

The bivariate Pareto copula

Question 3

Deduce the copula density function.

The bivariate Pareto copula

We have:

$$\begin{aligned}
\frac{\partial\,\mathbf{C}(u_1,u_2)}{\partial\,u_1} &= 1-\alpha\left(\left(1-u_1\right)^{-1/\alpha}+\left(1-u_2\right)^{-1/\alpha}-1\right)^{-\alpha-1}\times \\
&\quad \left(-\frac{1}{\alpha}\right)\left(1-u_1\right)^{-1/\alpha-1}\times(-1) \\
&= 1-\left(\left(1-u_1\right)^{-1/\alpha}+\left(1-u_2\right)^{-1/\alpha}-1\right)^{-\alpha-1}\times \\
&\quad \left(1-u_1\right)^{-1/\alpha-1}
\end{aligned}$$

The bivariate Pareto copula

We deduce that the probability density function of the copula is:

$$\begin{aligned}
c(u_1, u_2) &= \frac{\partial^2 \mathbf{C}(u_1, u_2)}{\partial\, u_1\, \partial\, u_2} \\
&= -(-\alpha - 1)\left((1-u_1)^{-1/\alpha} + (1-u_2)^{-1/\alpha} - 1\right)^{-\alpha-2} \times \\
&\quad \left(-\frac{1}{\alpha}\right)(1-u_2)^{-1/\alpha-1} \times (-1) \times (1-u_1)^{-1/\alpha-1} \\
&= \left(\frac{\alpha+1}{\alpha}\right)\left((1-u_1)^{-1/\alpha} + (1-u_2)^{-1/\alpha} - 1\right)^{-\alpha-2} \times \\
&\quad (1 - u_1 - u_2 + u_1 u_2)^{-1/\alpha-1}
\end{aligned}$$

The bivariate Pareto copula

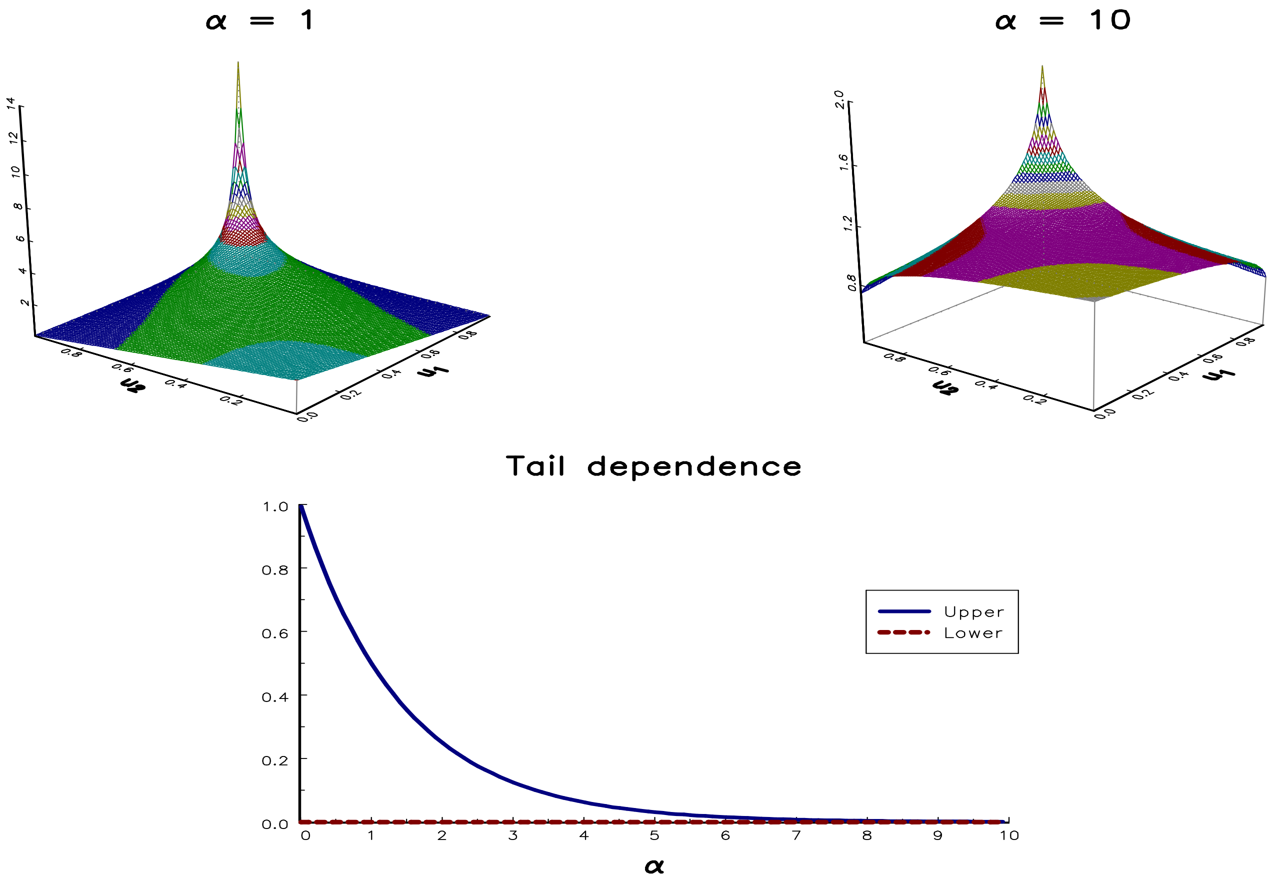
Remark

Another expression of $c(u_1, u_2)$ is:

$$\begin{aligned} c(u_1, u_2) &= \left(\frac{\alpha+1}{\alpha}\right)\left((1-u_1)(1-u_2)\right)^{1/\alpha} \times \\ &\quad \left((1-u_1)^{1/\alpha} + (1-u_2)^{1/\alpha} - (1-u_1)^{1/\alpha}(1-u_2)^{1/\alpha}\right)^{-\alpha-2} \end{aligned}$$

The bivariate Pareto copula

In this Figure, we have reported the density of the Pareto copula when α is equal to 1 and 10.



The bivariate Pareto copula

Question 4

Show that the bivariate Pareto copula function has no lower tail dependence, but an upper tail dependence.

The bivariate Pareto copula

We have:

$$\begin{aligned}
\lambda^- &= \lim_{u\to 0^+} \frac{\mathbf{C}(u,u)}{u} \\
&= 2 \lim_{u\to 0^+} \frac{\partial\, \mathbf{C}(u,u)}{\partial\, u_1} \\
&= 2 \lim_{u\to 0^+} 1 - \left((1-u)^{-1/\alpha} + (1-u)^{-1/\alpha} - 1\right)^{-\alpha-1} (1-u)^{-1/\alpha-1} \\
&= 2 \lim_{u\to 0^+} (1-1) \\
&= 0
\end{aligned}$$

The bivariate Pareto copula

We have:

$$
\begin{aligned}
\lambda^{+} &= \lim_{u \to 1^{-}} \frac{1 - 2u + \mathbf{C}(u, u)}{1 - u} \\
&= \lim_{u \to 1^{-}} \frac{\left((1-u)^{-1/\alpha} + (1-u)^{-1/\alpha} - 1 \right)^{-\alpha}}{1 - u} \\
&= \lim_{u \to 1^{-}} \left(1 + 1 - (1-u)^{1/\alpha} \right)^{-\alpha} \\
&= 2^{-\alpha}
\end{aligned}
$$

The tail dependence coefficients λ^{-} and λ^{+} are given with respect to the parameter α in previous Figure. We deduce that the bivariate Pareto copula function has no lower tail dependence ($\lambda^{-} = 0$), but an upper tail dependence ($\lambda^{+} = 2^{-\alpha}$).

The bivariate Pareto copula

Question 5

Do you think that the bivariate Pareto copula family can reach the copula functions $\mathbf{C}^{-}$, $\mathbf{C}^{\perp}$ and $\mathbf{C}^{+}$? Justify your answer.

The bivariate Pareto copula

The bivariate Pareto copula family cannot reach $\mathbf{C}^-$ because λ^- is never equal to 1. We notice that:

$$\lim_{\alpha \to \infty} \lambda^+ = 0$$

and

$$\lim_{\alpha \to 0} \lambda^+ = 1$$

This implies that the bivariate Pareto copula may reach $\mathbf{C}^\perp$ and $\mathbf{C}^+$ for these two limit cases: $\alpha \to \infty$ and $\alpha \to 0$. In fact, $\alpha \to 0$ does not correspond to the copula $\mathbf{C}^+$ because λ^- is always equal to 0.

The bivariate Pareto copula

Question 6

Let X_1 and X_2 be two Pareto-distributed random variables, whose parameters are (α_1, θ_1) and (α_2, θ_2).

The bivariate Pareto copula

Question 6.a

Show that the linear correlation between X_1 and X_2 is equal to 1 if and only if the parameters α_1 and α_2 are equal.

The bivariate Pareto copula

We note $U_1 = \mathbf{F}_1(X_1)$ and $U_2 = \mathbf{F}_2(X_2)$. X_1 and X_2 are comonotonic if and only if:

$$U_2 = U_1$$

We deduce that:

$$\begin{aligned}
1 - \left(\frac{\theta_2 + X_2}{\theta_2}\right)^{-\alpha_2} &= 1 - \left(\frac{\theta_1 + X_1}{\theta_1}\right)^{-\alpha_1} \\
\Leftrightarrow \quad \left(\frac{\theta_2 + X_2}{\theta_2}\right)^{-\alpha_2} &= \left(\frac{\theta_1 + X_1}{\theta_1}\right)^{-\alpha_1} \\
\Leftrightarrow \quad X_2 &= \theta_2 \left(\left(\frac{\theta_1 + X_1}{\theta_1}\right)^{\alpha_1/\alpha_2} - 1\right)
\end{aligned}$$

We know that $\rho\left\langle X_1, X_2\right\rangle = 1$ if and only if there is an increasing linear relationship between X_1 and X_2. This implies that:

$$\frac{\alpha_1}{\alpha_2} = 1$$

The bivariate Pareto copula

Question 6.b

Show that the linear correlation between X_1 and X_2 can never reached the lower bound -1.

The bivariate Pareto copula

X_1 and X_2 are countermonotonic if and only if:

$$U_2 = 1 - U_1$$

We deduce that:

$$\begin{aligned} & \left(\frac{\theta_2 + X_2}{\theta_2}\right)^{-\alpha_2} = 1 - \left(\frac{\theta_1 + X_1}{\theta_1}\right)^{-\alpha_1} \\ \Leftrightarrow \quad & \left(\frac{\theta_2 + X_2}{\theta_2}\right)^{-\alpha_2} = 1 - \left(\frac{\theta_1 + X_1}{\theta_1}\right)^{-\alpha_1} \\ \Leftrightarrow \quad & X_2 = \theta_2 \left(\left(1 - \left(\frac{\theta_1 + X_1}{\theta_1}\right)^{-\alpha_1}\right)^{1/\alpha_2} - 1\right) \end{aligned}$$

It is not possible to obtain a decreasing linear function between X_1 and X_2. This implies that $\rho \langle X_1, X_2 \rangle > -1$.

The bivariate Pareto copula

Question 6.c

Build a new bivariate Pareto distribution by assuming that the marginal distributions are $\mathcal{P}(\alpha_1, \theta_1)$ and $\mathcal{P}(\alpha_2, \theta_2)$ and the dependence is a bivariate Pareto copula function with parameter α. What is the relevance of this approach for building bivariate Pareto distributions?

The bivariate Pareto copula

We have:

$$\begin{aligned}\mathbf{F}'(x_1, x_2) &= \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \\ &= 1 - \left(\frac{\theta_1 + x_1}{\theta_1}\right)^{-\alpha_1} - \left(\frac{\theta_2 + x_2}{\theta_2}\right)^{-\alpha_2} + \\ &\quad \left(\left(\frac{\theta_1 + x_1}{\theta_1}\right)^{\alpha_1/\alpha} + \left(\frac{\theta_2 + x_2}{\theta_2}\right)^{\alpha_2/\alpha} - 1\right)^{-\alpha}\end{aligned}$$

The traditional bivariate Pareto distribution $\mathbf{F}(x_1, x_2)$ is a special case of $\mathbf{F}'(x_1, x_2)$ when:

$$\alpha_1 = \alpha_2 = \alpha$$

Using $\mathbf{F}'$ instead of $\mathbf{F}$, we can control the tail dependence, but also the univariate tail index of the two margins.

Calculation of correlation bounds

Question 1

Give the mathematical definition of the copula functions $\mathbf{C}^{-}$, $\mathbf{C}^{\perp}$ and $\mathbf{C}^{+}$. What is the probabilistic interpretation of these copulas?

Calculation of correlation bounds

We have:

$$\begin{aligned}
\mathbf{C}^{-}(u_1,u_2) &= \max(u_1+u_2-1,0) \\
\mathbf{C}^{\perp}(u_1,u_2) &= u_1 u_2 \\
\mathbf{C}^{+}(u_1,u_2) &= \min(u_1,u_2)
\end{aligned}$$

Let X_1 and X_2 be two random variables. We have:

(i) $\mathbf{C}\langle X_1,X_2\rangle = \mathbf{C}^{-}$ if and only if there exists a non-increasing function f such that we have $X_2 = f(X_1)$;

(ii) $\mathbf{C}\langle X_1,X_2\rangle = \mathbf{C}^{\perp}$ if and only if X_1 and X_2 are independent;

(iii) $\mathbf{C}\langle X_1,X_2\rangle = \mathbf{C}^{+}$ if and only if there exists a non-decreasing function f such that we have $X_2 = f(X_1)$.

Calculation of correlation bounds

Question 2

We note $\boldsymbol{\tau}$ and LGD the default time and the loss given default of a counterparty. We assume that $\boldsymbol{\tau} \sim \mathcal{E}(\lambda)$ and $\mathrm{LGD} \sim \mathcal{U}_{[0,1]}$.

Calculation of correlation bounds

We note $U_1 = 1 - \exp(-\lambda \boldsymbol{\tau})$ and $U_2 = \mathrm{LGD}$.

Calculation of correlation bounds

Question 2.a

Show that the dependence between $\boldsymbol{\tau}$ and LGD is maximum when the following equality holds:

$$\mathrm{LGD} + e^{-\lambda \boldsymbol{\tau}} - 1 = 0$$

Calculation of correlation bounds

The dependence between $\boldsymbol{\tau}$ and LGD is maximum when we have $\mathbf{C}\left\langle \boldsymbol{\tau}, \mathrm{LGD} \right\rangle = \mathbf{C}^{+}$. Since we have $U_1 = U_2$, we conclude that:

$$\mathrm{LGD} + e^{-\lambda \boldsymbol{\tau}} - 1 = 0$$

Calculation of correlation bounds

Question 2.b

Show that the linear correlation $\rho\left(\boldsymbol{\tau}, \mathrm{LGD}\right)$ verifies the following inequality:

$$\left|\rho\left\langle\boldsymbol{\tau}, \mathrm{LGD}\right\rangle\right| \leq \frac{\sqrt{3}}{2}$$

Calculation of correlation bounds

We know that:

$$\rho \langle \boldsymbol{\tau}, \mathrm{LGD} \rangle \in [\rho_{\min} \langle \boldsymbol{\tau}, \mathrm{LGD} \rangle, \rho_{\max} \langle \boldsymbol{\tau}, \mathrm{LGD} \rangle]$$

where $\rho_{\min} \langle \boldsymbol{\tau}, \mathrm{LGD} \rangle$ (resp. $\rho_{\max} \langle \boldsymbol{\tau}, \mathrm{LGD} \rangle$) is the linear correlation corresponding to the copula $\mathbf{C}^-$ (resp. $\mathbf{C}^+$). It comes that:

$$\mathbb{E}[\boldsymbol{\tau}] = \sigma(\boldsymbol{\tau}) = \frac{1}{\lambda}$$

and:

$$\begin{aligned} \mathbb{E}[\mathrm{LGD}] &= \frac{1}{2} \\ \sigma(\mathrm{LGD}) &= \sqrt{\frac{1}{12}} \end{aligned}$$

Calculation of correlation bounds

In the case $\mathbf{C}\langle\boldsymbol{\tau},\mathrm{LGD}\rangle = \mathbf{C}^-$, we have $U_1 = 1 - U_2$. It follows that $\mathrm{LGD} = e^{-\lambda\boldsymbol{\tau}}$. We have:

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\tau}\,\mathrm{LGD}\right] = \mathbb{E}\left[\boldsymbol{\tau} e^{-\lambda\boldsymbol{\tau}}\right] &= \int_0^\infty t e^{-\lambda t}\lambda e^{-\lambda t}\,\mathrm{d}t \\
&= \int_0^\infty t\lambda e^{-2\lambda t}\,\mathrm{d}t \\
&= \left[-\frac{t e^{-2\lambda t}}{2}\right]_0^\infty + \frac{1}{2}\int_0^\infty e^{-2\lambda t}\,\mathrm{d}t \\
&= 0 + \frac{1}{2}\left[-\frac{e^{-2\lambda t}}{2\lambda}\right]_0^\infty \\
&= \frac{1}{4\lambda}
\end{aligned}$$

We deduce that:

$$\rho_{\min}\langle\boldsymbol{\tau},\mathrm{LGD}\rangle = \left(\frac{1}{4\lambda} - \frac{1}{2\lambda}\right) \Big/ \left(\frac{1}{\lambda}\sqrt{\frac{1}{12}}\right) = -\frac{\sqrt{3}}{2}$$

Calculation of correlation bounds

In the case $\mathbf{C}\langle\boldsymbol{\tau},\mathrm{LGD}\rangle=\mathbf{C}^{+}$, we have $\mathrm{LGD}=1-e^{-\lambda\tau}$. We have:

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\tau}\,\mathrm{LGD}\right]=\mathbb{E}\left[\boldsymbol{\tau}\left(1-e^{-\lambda\boldsymbol{\tau}}\right)\right] &= \int_0^\infty t\left(1-e^{-\lambda t}\right)\lambda e^{-\lambda t}\,\mathrm{d}t \\
&= \int_0^\infty t\lambda e^{-\lambda t}\,\mathrm{d}t-\int_0^\infty t\lambda e^{-2\lambda t}\,\mathrm{d}t \\
&= \left(\left[-te^{-\lambda t}\right]_0^\infty+\int_0^\infty e^{-\lambda t}\,\mathrm{d}t\right)-\frac{1}{4\lambda} \\
&= 0+\left[-\frac{e^{-\lambda t}}{\lambda}\right]_0^\infty-\frac{1}{4\lambda} \\
&= \frac{3}{4\lambda}
\end{aligned}$$

We deduce that:

$$\rho_{\max}\langle\boldsymbol{\tau},\mathrm{LGD}\rangle=\left(\frac{3}{4\lambda}-\frac{1}{2\lambda}\right)\Big/\left(\frac{1}{\lambda}\sqrt{\frac{1}{12}}\right)=\frac{\sqrt{3}}{2}$$

Calculation of correlation bounds

We finally obtain the following result:

$$\left|\rho\left\langle \boldsymbol{\tau}, \mathrm{LGD}\right\rangle\right| \leq \frac{\sqrt{3}}{2}$$

Calculation of correlation bounds

Question 2.c

Comment on these results.

Calculation of correlation bounds

We notice that $|\rho\langle\boldsymbol{\tau},\mathrm{LGD}\rangle|$ is lower than 86.6%, implying that the bounds -1 and $+1$ can not be reached.

Calculation of correlation bounds

Question 3

We consider two exponential default times $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ with parameters λ_1 and λ_2.

Calculation of correlation bounds

Question 3.a

We assume that the dependence function between $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ is $\mathbf{C}^+$. Demonstrate that the following relation is true:

$$\boldsymbol{\tau}_1 = \frac{\lambda_2}{\lambda_1}\boldsymbol{\tau}_2$$

Calculation of correlation bounds

If the copula function of $(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$ is the Fréchet upper bound copula, $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ are comonotone. We deduce that:

$$U_1 = U_2 \Longleftrightarrow 1 - e^{-\lambda_1 \boldsymbol{\tau}_1} = 1 - e^{-\lambda_2 \boldsymbol{\tau}_2}$$

and:

$$\boldsymbol{\tau}_1 = \frac{\lambda_2}{\lambda_1} \boldsymbol{\tau}_2$$

Calculation of correlation bounds

Question 3.b

Show that there exists a function f such that $\boldsymbol{\tau}_2 = f\left(\boldsymbol{\tau}_2\right)$ when the dependence function is $\mathbf{C}^-$.

Calculation of correlation bounds

We have $U_1 = 1 - U_2$. It follows that $\mathbf{S}_1(\boldsymbol{\tau}_1) = 1 - \mathbf{S}_2(\boldsymbol{\tau}_2)$. We deduce that:

$$e^{-\lambda_1 \boldsymbol{\tau}_1} = 1 - e^{-\lambda_2 \boldsymbol{\tau}_2}$$

and:

$$\boldsymbol{\tau}_1 = \frac{-\ln\left(1 - e^{-\lambda_2 \boldsymbol{\tau}_2}\right)}{\lambda_1}$$

There exists then a function f such that $\boldsymbol{\tau}_1 = f(\boldsymbol{\tau}_2)$ with:

$$f(t) = \frac{-\ln\left(1 - e^{-\lambda_2 t}\right)}{\lambda_1}$$

Calculation of correlation bounds

Question 3.c

Show that the lower and upper bounds of the linear correlation satisfy the following relationship:

$$-1 < \rho \langle \boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \rangle \leq 1$$

Calculation of correlation bounds

Using Question 2(b), we known that $\rho \in [\rho_{\min}, \rho_{\max}]$ where $\rho_{\min}$ and $\rho_{\max}$ are the correlations of $(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$ when the copula function is respectively $\mathbf{C}^-$ and $\mathbf{C}^+$. We also know that $\rho = 1$ (resp. $\rho = -1$) if there exists a linear and increasing (resp. decreasing) function f such that $\boldsymbol{\tau}_1 = f(\boldsymbol{\tau}_2)$. When the copula is $\mathbf{C}^+$, we have $f(t) = \frac{\lambda_2}{\lambda_1} t$ and $f'(t) = \frac{\lambda_2}{\lambda_1} > 0$. As it is a linear and increasing function, we deduce that $\rho_{\max} = 1$. When the copula is $\mathbf{C}^-$, we have:

$$f(t) = \frac{-\ln\left(1 - e^{-\lambda_2 t}\right)}{\lambda_1}$$

and:

$$f'(t) = -\frac{\lambda_2 e^{-\lambda_2 t} \ln\left(1 - e^{-\lambda_2 t}\right)}{\lambda_1 \left(1 - e^{-\lambda_2 t}\right)} < 0$$

The function $f(t)$ is decreasing, but it is not linear. We deduce that $\rho_{\min} \neq -1$ and:

$$-1 < \rho \leq 1$$

Calculation of correlation bounds

Question 3.d

In the more general case, show that the linear correlation of a random vector (X_1, X_2) can not be equal to -1 if the support of the random variables X_1 and X_2 is $[0, +\infty]$.

Calculation of correlation bounds

When the copula is $\mathbf{C}^{-}$, we know that there exists a decreasing function f such that $X_2 = f(X_1)$. We also know that the linear correlation reaches the lower bound -1 if the function f is linear:

$$X_2 = a + bX_1$$

This implies that $b < 0$. When X_1 takes the value $+\infty$, we obtain:

$$X_2 = a + b \times \infty$$

As the lower bound of X_2 is equal to zero 0, we deduce that $a = +\infty$. This means that the function $f(x) = a + bx$ does not exist. We conclude that the lower bound $\rho = -1$ can not be reached.

Calculation of correlation bounds

Question 4

We assume that (X_1, X_2) is a Gaussian random vector where $X_1 \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right)$, $X_2 \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right)$ and ρ is the linear correlation between X_1 and X_2. We note $\theta = (\mu_1, \sigma_1, \mu_2, \sigma_2, \rho)$ the set of parameters.

Calculation of correlation bounds

Question 4.a

Find the probability distribution of $X_1 + X_2$.

Calculation of correlation bounds

$X_1 + X_2$ is a Gaussian random variable because it is a linear combination of the Gaussian random vector (X_1, X_2). We have:

$$\mathbb{E}\left[X_1 + X_2\right] = \mu_1 + \mu_2$$

and:

$$\operatorname{var}\left(X_1 + X_2\right) = \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2$$

We deduce that:

$$X_1 + X_2 \sim \mathcal{N}\left(\mu_1 + \mu_2, \sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2\right)$$

Calculation of correlation bounds

Question 4.b

Then show that the covariance between $Y_1 = e^{X_1}$ and $Y_2 = e^{X_2}$ is equal to:

$$\operatorname{cov}(Y_1, Y_2) = e^{\mu_1 + \frac{1}{2}\sigma_1^2} e^{\mu_2 + \frac{1}{2}\sigma_2^2} \left(e^{\rho\sigma_1\sigma_2} - 1\right)$$

Calculation of correlation bounds

We have:

$$\begin{aligned}\operatorname{cov}(Y_1, Y_2) &= \mathbb{E}[Y_1 Y_2] - \mathbb{E}[Y_2]\,\mathbb{E}[Y_2] \\ &= \mathbb{E}\left[e^{X_1+X_2}\right] - \mathbb{E}[Y_2]\,\mathbb{E}[Y_2]\end{aligned}$$

We know that $e^{X_1+X_2}$ is a lognormal random variable. We deduce that:

$$\begin{aligned}\mathbb{E}\left[e^{X_1+X_2}\right] &= \exp\left(\mathbb{E}[X_1+X_2] + \frac{1}{2}\operatorname{var}(X_1+X_2)\right) \\ &= \exp\left(\mu_1+\mu_2+\frac{1}{2}\left(\sigma_1^2+2\rho\sigma_1\sigma_2+\sigma_2^2\right)\right) \\ &= e^{\mu_1+\frac{1}{2}\sigma_1^2}e^{\mu_2+\frac{1}{2}\sigma_2^2}e^{\rho\sigma_1\sigma_2}\end{aligned}$$

We finally obtain:

$$\operatorname{cov}(Y_1, Y_2) = e^{\mu_1+\frac{1}{2}\sigma_1^2}e^{\mu_2+\frac{1}{2}\sigma_2^2}\left(e^{\rho\sigma_1\sigma_2}-1\right)$$

Calculation of correlation bounds

Question 4.c

Deduce the correlation between Y_1 and Y_2.

Calculation of correlation bounds

We have:

$$\begin{aligned}
\rho\left\langle Y_1, Y_2\right\rangle &= \frac{e^{\mu_1+\frac{1}{2}\sigma_1^2}e^{\mu_2+\frac{1}{2}\sigma_2^2}\left(e^{\rho\sigma_1\sigma_2}-1\right)}{\sqrt{e^{2\mu_1+\sigma_1^2}\left(e^{\sigma_1^2}-1\right)}\sqrt{e^{2\mu_2+\sigma_2^2}\left(e^{\sigma_2^2}-1\right)}} \\
&= \frac{e^{\rho\sigma_1\sigma_2}-1}{\sqrt{e^{\sigma_1^2}-1}\sqrt{e^{\sigma_2^2}-1}}
\end{aligned}$$

Calculation of correlation bounds

Question 4.d

For which values of θ does the equality $\rho\langle Y_1, Y_2\rangle = +1$ hold? Same question when $\rho\langle Y_1, Y_2\rangle = -1$.

Calculation of correlation bounds

$\rho\langle Y_1, Y_2\rangle$ is an increasing function with respect to ρ. We deduce that:

$$\rho\langle Y_1, Y_2\rangle = 1 \Longleftrightarrow \rho = 1 \text{ and } \sigma_1 = \sigma_2$$

The lower bound of $\rho\langle Y_1, Y_2\rangle$ is reached if ρ is equal to -1. In this case, we have:

$$\rho\langle Y_1, Y_2\rangle = \frac{e^{-\sigma_1\sigma_2} - 1}{\sqrt{e^{\sigma_1^2} - 1}\sqrt{e^{\sigma_2^2} - 1}} > -1$$

It follows that $\rho\langle Y_1, Y_2\rangle \neq -1$.

Calculation of correlation bounds

Question 4.e

We consider the bivariate Black-Scholes model:

$$\begin{cases} \mathrm{d}S_1(t) = \mu_1 S_1(t)\,\mathrm{d}t + \sigma_1 S_1(t)\,\mathrm{d}W_1(t) \\ \mathrm{d}S_2(t) = \mu_2 S_2(t)\,\mathrm{d}t + \sigma_2 S_2(t)\,\mathrm{d}W_2(t) \end{cases}$$

with $\mathbb{E}[W_1(t) W_2(t)] = \rho t$. Deduce the linear correlation between $S_1(t)$ and $S_2(t)$. Find the limit case $\lim_{t \to \infty} \rho \langle S_1(t), S_2(t) \rangle$.

Calculation of correlation bounds

It is obvious that:

$$\rho \left\langle S_1(t), S_2(t) \right\rangle = \frac{e^{\rho \sigma_1 \sigma_2 t} - 1}{\sqrt{e^{\sigma_1^2 t} - 1}\sqrt{e^{\sigma_2^2 t} - 1}}$$

In the case $\sigma_1 = \sigma_2$ and $\rho = 1$, we have $\rho \left\langle S_1(t), S_2(t) \right\rangle = 1$. Otherwise, we obtain:

$$\lim_{t \to \infty} \rho \left\langle S_1(t), S_2(t) \right\rangle = 0$$

Calculation of correlation bounds

Question 4.f

Comment on these results.

Calculation of correlation bounds

In the case of lognormal random variables, the linear correlation does not necessarily range between -1 and $+1$.

Extreme value theory in the bivariate case

Question 1

What is an extreme value (EV) copula **C**?

Extreme value theory in the bivariate case

An extreme value copula **C** satisfies the following relationship:

$$\mathbf{C}\left(u_1^t, u_2^t\right) = \mathbf{C}^t\left(u_1, u_2\right)$$

for all $t > 0$.

Extreme value theory in the bivariate case

Question 2

Show that $\mathbf{C}^{\perp}$ and $\mathbf{C}^{+}$ are EV copulas. Why $\mathbf{C}^{-}$ can not be an EV copula?

Extreme value theory in the bivariate case

The product copula $\mathbf{C}^{\perp}$ is an EV copula because we have:

$$\begin{aligned} \mathbf{C}^{\perp}\left(u_1^t, u_2^t\right) &= u_1^t u_2^t \\ &= \left(u_1 u_2\right)^t \\ &= \left[\mathbf{C}^{\perp}\left(u_1, u_2\right)\right]^t \end{aligned}$$

Extreme value theory in the bivariate case

For the copula $\mathbf{C}^{+}$, we obtain:

$$\begin{aligned}
\mathbf{C}^{+}\left(u_1^t, u_2^t\right) &= \min\left(u_1^t, u_2^t\right) \\
&= \begin{cases} u_1^t & \text{if } u_1 \leq u_2 \\ u_2^t & \text{otherwise} \end{cases} \\
&= \left(\min\left(u_1, u_2\right)\right)^t \\
&= \left[\mathbf{C}^{+}\left(u_1, u_2\right)\right]^t
\end{aligned}$$

Extreme value theory in the bivariate case

However, the EV property does not hold for the Fréchet lower bound copula $\mathbf{C}^-$:

$$\mathbf{C}^-\left(u_1^t, u_2^t\right) = \max\left(u_1^t + u_2^t - 1, 0\right) \neq \max\left(u_1 + u_2 - 1, 0\right)^t$$

Indeed, we have $\mathbf{C}^-(0.5, 0.8) = \max(0.5 + 0.8 - 1, 0) = 0.3$ and:

$$\begin{array}{rcl}\mathbf{C}^-\left(0.5^2, 0.8^2\right) & = & \max(0.25 + 0.64 - 1, 0) \\ & = & 0 \\ & \neq & 0.3^2\end{array}$$

Extreme value theory in the bivariate case

Question 3

We define the Gumbel-Hougaard copula as follows:

$$\mathbf{C}(u_1, u_2) = \exp\left(-\left[(-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}\right]^{1/\theta}\right)$$

with $\theta \geq 1$. Verify that it is an EV copula.

Extreme value theory in the bivariate case

We have:

$$\begin{aligned}
\mathbf{C}\left(u_1^t, u_2^t\right) &= \exp\left(-\left[\left(-\ln u_1^t\right)^\theta + \left(-\ln u_2^t\right)^\theta\right]^{1/\theta}\right) \\
&= \exp\left(-\left[\left(-t\ln u_1\right)^\theta + \left(-t\ln u_2\right)^\theta\right]^{1/\theta}\right) \\
&= \exp\left(-t\left[\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right]^{1/\theta}\right) \\
&= \left(e^{-\left[\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right]^{1/\theta}}\right)^t \\
&= \mathbf{C}^t\left(u_1, u_2\right)
\end{aligned}$$

Extreme value theory in the bivariate case

Question 4

What is the definition of the upper tail dependence λ? What is its usefulness in multivariate extreme value theory?

Extreme value theory in the bivariate case

The upper tail dependence λ is defined as follows:

$$\lambda = \lim_{u \to 1^{+}} \frac{1 - 2u + \mathbf{C}\left(u_1, u_2\right)}{1 - u}$$

It measures the probability to have an extreme in one direction knowing that we have already an extreme in the other direction. If λ is equal to 0, extremes are independent and the EV copula is the product copula $\mathbf{C}^{\perp}$. If λ is equal to 1, extremes are comonotonic and the EV copula is the Fréchet upper bound copula $\mathbf{C}^{+}$. Moreover, the upper tail dependence of the copula between the random variables is equal to the upper tail dependence of the copula between the extremes.

Extreme value theory in the bivariate case

Question 5

Let $f(x)$ and $g(x)$ be two functions such that $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$. If $g'(x_0) \neq 0$, L'Hospital's rule states that:

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$$

Deduce that the upper tail dependence λ of the Gumbel-Hougaard copula is $2 - 2^{1/\theta}$. What is the correlation of two extremes when $\theta = 1$?

Extreme value theory in the bivariate case

Using L'Hospital's rule, we have:

$$\begin{aligned}
\lambda &= \lim_{u \to 1^+} \frac{1 - 2u + e^{-\left[(-\ln u)^\theta + (-\ln u)^\theta\right]^{1/\theta}}}{1-u} \\
&= \lim_{u \to 1^+} \frac{1 - 2u + e^{-\left[2(-\ln u)^\theta\right]^{1/\theta}}}{1-u} \\
&= \lim_{u \to 1^+} \frac{1 - 2u + u^{2^{1/\theta}}}{1-u} \\
&= \lim_{u \to 1^+} \frac{0 - 2 + 2^{1/\theta} u^{2^{1/\theta}-1}}{-1} \\
&= \lim_{u \to 1^+} 2 - 2^{1/\theta} u^{2^{1/\theta}-1} \\
&= 2 - 2^{1/\theta}
\end{aligned}$$

Extreme value theory in the bivariate case

If θ is equal to 1, we obtain $\lambda = 0$. It comes that the EV copula is the product copula. Extremes are then not correlated. This result is not surprising because the Gumbel-Houggard copula is equal to the product copula when $\theta = 1$:

$$e^{-\left[(-\ln u_1)^1 + (-\ln u_2)^1\right]^1} = u_1 u_2 = \mathbf{C}^{\perp}(u_1, u_2)$$

Extreme value theory in the bivariate case

Question 6

We define the Marshall-Olkin copula as follows:

$$\mathbf{C}(u_1, u_2) = u_1^{1-\theta_1} u_2^{1-\theta_2} \min\left(u_1^{\theta_1}, u_2^{\theta_2}\right)$$

with $\{\theta_1, \theta_2\} \in [0, 1]^2$.

Extreme value theory in the bivariate case

Question 6.a

Verify that it is an EV copula.

Extreme value theory in the bivariate case

We have:

$$\begin{aligned}
\mathbf{C}\left(u_1^t, u_2^t\right) &= u_1^{t(1-\theta_1)} u_2^{t(1-\theta_2)} \min\left(u_1^{t\theta_1}, u_2^{t\theta_2}\right) \\
&= \left(u_1^{1-\theta_1}\right)^t \left(u_2^{1-\theta_2}\right)^t \left(\min\left(u_1^{\theta_1}, u_2^{\theta_2}\right)\right)^t \\
&= \left(u_1^{1-\theta_1} u_2^{1-\theta_2} \min\left(u_1^{\theta_1}, u_2^{\theta_2}\right)\right)^t \\
&= \mathbf{C}^t\left(u_1, u_2\right)
\end{aligned}$$

Extreme value theory in the bivariate case

Question 6.b

Find the upper tail dependence λ of the Marshall-Olkin copula.

Extreme value theory in the bivariate case

If $\theta_1 > \theta_2$, we obtain:

$$\begin{aligned}
\lambda &= \lim_{u \to 1^+} \frac{1 - 2u + u^{1-\theta_1} u^{1-\theta_2} \min\left(u^{\theta_1}, u^{\theta_2}\right)}{1 - u} \\
&= \lim_{u \to 1^+} \frac{1 - 2u + u^{1-\theta_1} u^{1-\theta_2} u^{\theta_1}}{1 - u} \\
&= \lim_{u \to 1^+} \frac{1 - 2u + u^{2-\theta_2}}{1 - u} \\
&= \lim_{u \to 1^+} \frac{0 - 2 + (2 - \theta_2) u^{1-\theta_2}}{-1} \\
&= \lim_{u \to 1^+} 2 - 2u^{1-\theta_2} + \theta_2 u^{1-\theta_2} \\
&= \theta_2
\end{aligned}$$

If $\theta_2 > \theta_1$, we have $\lambda = \theta_1$. We deduce that the upper tail dependence of the Marshall-Olkin copula is $\min(\theta_1, \theta_2)$.

Extreme value theory in the bivariate case

Question 6.c

What is the correlation of two extremes when $\min(\theta_1, \theta_2) = 0$?

Extreme value theory in the bivariate case

If $\theta_1 = 0$ or $\theta_2 = 0$, we obtain $\lambda = 0$. It comes that the copula of the extremes is the product copula. Extremes are then not correlated.

Extreme value theory in the bivariate case

Question 6.d

In which case are two extremes perfectly correlated?

Extreme value theory in the bivariate case

Two extremes are perfectly correlated when we have $\theta_1 = \theta_2 = 1$. In this case, we obtain:

$$\mathbf{C}(u_1, u_2) = \min(u_1, u_2) = \mathbf{C}^{+}(u_1, u_2)$$

Maximum domain of attraction in the bivariate case

Question 1

We consider the following distributions of probability:

Distribution		$\mathbf{F}(x)$
Exponential	$\mathcal{E}(\lambda)$	$1 - e^{-\lambda x}$
Uniform	$\mathcal{U}_{[0,1]}$	x
Pareto	$\mathcal{P}(\alpha, \theta)$	$1 - \left(\frac{\theta + x}{\theta}\right)^{-\alpha}$

Maximum domain of attraction in the bivariate case

Question 1

For each distribution, we give the normalization parameters a_n and b_n of the Fisher-Tippet theorem and the corresponding limit distribution distribution $\mathbf{G}(x)$:

Distribution	a_n	b_n	$\mathbf{G}(x)$
Exponential	λ^{-1}	$\lambda^{-1}\ln n$	$\mathbf{\Lambda}(x) = e^{-e^{-x}}$
Uniform	n^{-1}	$1 - n^{-1}$	$\mathbf{\Psi}_1(x-1) = e^{x-1}$
Pareto	$\theta\alpha^{-1}n^{1/\alpha}$	$\theta n^{1/\alpha} - \theta$	$\mathbf{\Phi}_\alpha\left(1 + \frac{x}{\alpha}\right) = e^{-\left(1+\frac{x}{\alpha}\right)^{-\alpha}}$

We note $\mathbf{G}(x_1, x_2)$ the asymptotic distribution of the bivariate random vector $(X_{1,n:n}, X_{2,n:n})$ where $X_{1,i}$ (resp. $X_{2,i}$) are *iid* random variables.

Maximum domain of attraction in the bivariate case

Let (X_1, X_2) be a bivariate random variable whose probability distribution is:

$$\mathbf{F}(x_1, x_2) = \mathbf{C}_{\langle X_1, X_2 \rangle}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$$

We know that the corresponding EV probability distribution is:

$$\mathbf{G}(x_1, x_2) = \mathbf{C}^{\star}_{\langle X_1, X_2 \rangle}(\mathbf{G}_1(x_1), \mathbf{G}_2(x_2))$$

where $\mathbf{G}_1$ and $\mathbf{G}_2$ are the two univariate EV probability distributions and $\mathbf{C}^{\star}_{\langle X_1, X_2 \rangle}$ is the EV copula associated to $\mathbf{C}_{\langle X_1, X_2 \rangle}$.

Maximum domain of attraction in the bivariate case

Question 1.a

What is the expression of $\mathbf{G}(x_1, x_2)$ when $X_{1,i}$ and $X_{2,i}$ are independent, $X_{1,i} \sim \mathcal{E}(\lambda)$ and $X_{2,i} \sim \mathcal{U}_{[0,1]}$?

Maximum domain of attraction in the bivariate case

We deduce that:

$$\begin{aligned}
\mathbf{G}(x_1, x_2) &= \mathbf{C}^{\perp}(\mathbf{G}_1(x_1), \mathbf{G}_2(x_2)) \\
&= \mathbf{\Lambda}(x_1)\,\mathbf{\Psi}_1(x_2 - 1) \\
&= \exp\left(-e^{-x_1} + x_2 - 1\right)
\end{aligned}$$

Maximum domain of attraction in the bivariate case

Question 1.b

Same question when $X_{1,i} \sim \mathcal{E}(\lambda)$ and $X_{2,i} \sim \mathcal{P}(\theta, \alpha)$.

Maximum domain of attraction in the bivariate case

We have:

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= \mathbf{\Lambda}(x_1)\,\mathbf{\Phi}_\alpha\left(1 + \frac{x_2}{\alpha}\right) \\ &= \exp\left(-e^{-x_1} - \left(1 + \frac{x_2}{\alpha}\right)^{-\alpha}\right) \end{aligned}$$

Maximum domain of attraction in the bivariate case

Question 1.c

Same question when $X_{1,i} \sim \mathcal{U}_{[0,1]}$ and $X_{2,i} \sim \mathcal{P}(\theta, \alpha)$.

Maximum domain of attraction in the bivariate case

We have:

$$\begin{aligned}\mathbf{G}(x_1, x_2) &= \mathbf{\Psi}_1(x_1 - 1)\,\mathbf{\Phi}_\alpha\left(1 + \frac{x_2}{\alpha}\right) \\ &= \exp\left(x_1 - 1 - \left(1 + \frac{x_2}{\alpha}\right)^{-\alpha}\right)\end{aligned}$$

Maximum domain of attraction in the bivariate case

Question 2

What becomes the previous results when the dependence function between $X_{1,i}$ and $X_{2,i}$ is the Normal copula with parameter $\rho < 1$?

Maximum domain of attraction in the bivariate case

We know that the upper tail dependence is equal to zero for the Normal copula when $\rho < 1$. We deduce that the EV copula is the product copula. We then obtain the same results as previously.

Maximum domain of attraction in the bivariate case

Question 3

Same question when the parameter of the Normal copula is equal to one.

Maximum domain of attraction in the bivariate case

When the parameter ρ is equal to 1, the Normal copula is the Fréchet upper bound copula $\mathbf{C}^{+}$, which is an EV copula. We deduce the following results:

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= \min\left(\mathbf{\Lambda}(x_1), \mathbf{\Psi}_1(x_2 - 1)\right) \\ &= \min\left(\exp\left(-e^{-x_1}\right), \exp(x_2 - 1)\right) \end{aligned} \qquad \text{(a)}$$

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= \min\left(\mathbf{\Lambda}(x_1), \mathbf{\Phi}_\alpha\left(1 + \frac{x_2}{\alpha}\right)\right) \\ &= \min\left(\exp\left(-e^{-x_1}\right), \exp\left(-\left(1 + \frac{x_2}{\alpha}\right)^{-\alpha}\right)\right) \end{aligned} \qquad \text{(b)}$$

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= \min\left(\mathbf{\Psi}_1(x_1 - 1), \mathbf{\Phi}_\alpha\left(1 + \frac{x_2}{\alpha}\right)\right) \\ &= \min\left(\exp(x_2 - 1), \exp\left(-\left(1 + \frac{x_2}{\alpha}\right)^{-\alpha}\right)\right) \end{aligned} \qquad \text{(c)}$$

Maximum domain of attraction in the bivariate case

Question 4

Find the expression of $\mathbf{G}(x_1, x_2)$ when the dependence function is the Gumbel-Hougaard copula.

Maximum domain of attraction in the bivariate case

In the previous exercise, we have shown that the Gumbel-Houggard copula is an EV copula.

Maximum domain of attraction in the bivariate case

We deduce that:

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= e^{-\left[(-\ln \mathbf{\Lambda}(x_1))^{\theta} + (-\ln \mathbf{\Psi}_1(x_2 - 1))^{\theta}\right]^{1/\theta}} \\ &= \exp\left(-\left[e^{-\theta x_1} + (1 - x_2)^{\theta}\right]^{1/\theta}\right) \qquad \text{(a)} \end{aligned}$$

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= e^{-\left[(-\ln \mathbf{\Lambda}(x_1))^{\theta} + \left(-\ln \mathbf{\Phi}_{\alpha}\left(1 + \frac{x_2}{\alpha}\right)\right)^{\theta}\right]^{1/\theta}} \\ &= \exp\left(-\left[e^{-\theta x_1} + \left(1 + \frac{x_2}{\alpha}\right)^{-\alpha\theta}\right]^{1/\theta}\right) \qquad \text{(b)} \end{aligned}$$

$$\begin{aligned} \mathbf{G}(x_1, x_2) &= e^{-\left[(-\ln \mathbf{\Psi}_1(x_1 - 1))^{\theta} + \left(-\ln \mathbf{\Phi}_{\alpha}\left(1 + \frac{x_2}{\alpha}\right)\right)^{\theta}\right]^{1/\theta}} \\ &= \exp\left(-\left[(1 - x_1)^{\theta} + \left(1 + \frac{x_2}{\alpha}\right)^{-\alpha\theta}\right]^{1/\theta}\right) \qquad \text{(c)} \end{aligned}$$

Simulation of the bivariate Normal copula

Exercise

Let $X = (X_1, X_2)$ be a standard Gaussian vector with correlation ρ. We note $U_1 = \Phi(X_1)$ and $U_2 = \Phi(X_2)$.

Simulation of the bivariate Normal copula

Question 1

We note Σ the matrix defined as follows:

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Calculate the Cholesky decomposition of Σ. Deduce an algorithm to simulate X.

Simulation of the bivariate Normal copula

P is a lower triangular matrix such that we have $\Sigma = PP^\top$. We know that:

$$P = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$$

We verify that:

$$\begin{aligned} PP^\top &= \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} 1 & \rho \\ 0 & \sqrt{1-\rho^2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \end{aligned}$$

Simulation of the bivariate Normal copula

We deduce that:

$$\left(\begin{array}{c} X_1 \\ X_2 \end{array}\right) = \left(\begin{array}{cc} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{array}\right) \left(\begin{array}{c} N_1 \\ N_2 \end{array}\right)$$

where N_1 and N_2 are two independent standardized Gaussian random variables. Let n_1 and n_2 be two independent random variates, whose probability distribution is $\mathcal{N}(0,1)$. Using the Cholesky decomposition, we deduce that can simulate X in the following way:

$$\left\{ \begin{array}{l} x_1 \leftarrow n_1 \\ x_2 \leftarrow \rho n_1 + \sqrt{1-\rho^2} n_2 \end{array} \right.$$

Simulation of the bivariate Normal copula

Question 2

Show that the copula of (X_1, X_2) is the same that the copula of the random vector (U_1, U_2).

Simulation of the bivariate Normal copula

We have

$$\begin{aligned} \mathbf{C}\left\langle X_1, X_2 \right\rangle &= \mathbf{C}\left\langle \Phi\left(X_1\right), \Phi\left(X_2\right) \right\rangle \\ &= \mathbf{C}\left\langle U_1, U_2 \right\rangle \end{aligned}$$

because the function $\Phi(x)$ is non-decreasing. The copula of $U = (U_1, U_2)$ is then the copula of $X = (X_1, X_2)$.

Simulation of the bivariate Normal copula

Question 3

Deduce an algorithm to simulate the Normal copula with parameter ρ.

Simulation of the bivariate Normal copula

We deduce that we can simulate U with the following algorithm:

$$\left\{\begin{array}{l} u_1 \leftarrow \Phi\left(x_1\right) = \Phi\left(n_1\right) \\ u_2 \leftarrow \Phi\left(x_2\right) = \Phi\left(\rho n_1 + \sqrt{1-\rho^2} n_2\right) \end{array}\right.$$

Simulation of the bivariate Normal copula

Question 4

Calculate the conditional distribution of X_2 knowing that $X_1 = x$. Then show that:

$$\Phi_2\left(x_1, x_2; \rho\right) = \int_{-\infty}^{x_1} \Phi\left(\frac{x_2 - \rho x}{\sqrt{1-\rho^2}}\right) \phi\left(x\right) \mathrm{d}x$$

Simulation of the bivariate Normal copula

Let X_3 be a Gaussian random variable, which is independent from X_1 and X_2. Using the Cholesky decomposition, we know that:

$$X_2 = \rho X_1 + \sqrt{1-\rho^2} X_3$$

It follows that:

$$\begin{aligned}
\Pr\left\{ X_2 \leq x_2 \,\middle|\, X_1 = x \right\} &= \Pr\left\{ \rho X_1 + \sqrt{1-\rho^2} X_3 \leq x_2 \,\middle|\, X_1 = x \right\} \\
&= \Pr\left\{ X_3 \leq \frac{x_2 - \rho x}{\sqrt{1-\rho^2}} \right\} \\
&= \Phi\left(\frac{x_2 - \rho x}{\sqrt{1-\rho^2}} \right)
\end{aligned}$$

Simulation of the bivariate Normal copula

Then we deduce that:

$$\begin{aligned}
\Phi_2(x_1, x_2; \rho) &= \Pr\{X_1 \leq x_1, X_2 \leq x_2\} \\
&= \Pr\left\{X_1 \leq x_1, X_3 \leq \frac{x_2 - \rho X_1}{\sqrt{1-\rho^2}}\right\} \\
&= \mathbb{E}\left[\Pr\left\{X_1 \leq x_1, X_3 \leq \frac{x_2 - \rho X_1}{\sqrt{1-\rho^2}} \middle| X_1\right\}\right] \\
&= \int_{-\infty}^{x_1} \Phi\left(\frac{x_2 - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x)\,\mathrm{d}x
\end{aligned}$$

Simulation of the bivariate Normal copula

Question 5

Deduce an expression of the Normal copula.

Simulation of the bivariate Normal copula

Using the relationships $u_1 = \Phi(x_1)$, $u_2 = \Phi(x_2)$ and $\Phi_2(x_1, x_2; \rho) = \mathbf{C}(\Phi(x_1), \Phi(x_2); \rho)$, we obtain:

$$\begin{aligned}
\mathbf{C}(u_1, u_2; \rho) &= \int_{-\infty}^{\Phi^{-1}(u_1)} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x)\,\mathrm{d}x \\
&= \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho \Phi^{-1}(u)}{\sqrt{1-\rho^2}}\right) \mathrm{d}u
\end{aligned}$$

Simulation of the bivariate Normal copula

Question 6

Calculate the conditional copula function $\mathbf{C}_{2|1}$. Deduce an algorithm to simulate the Normal copula with parameter ρ.

Simulation of the bivariate Normal copula

We have:

$$\begin{aligned} \mathbf{C}_{2|1}(u_2 \mid u_1) &= \partial_{u_1} \mathbf{C}(u_1, u_2) \\ &= \Phi\left(\frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u_1)}{\sqrt{1-\rho^2}}\right) \end{aligned}$$

Let v_1 and v_2 be two independent uniform random variates. The simulation algorithm corresponds to the following steps:

$$\left\{ \begin{array}{l} u_1 = v_1 \\ \mathbf{C}_{2|1}(u_1, u_2) = v_2 \end{array} \right.$$

We deduce that:

$$\left\{ \begin{array}{l} u_1 \leftarrow v_1 \\ u_2 \leftarrow \Phi\left(\rho\Phi^{-1}(v_1) + \sqrt{1-\rho^2}\Phi^{-1}(v_2)\right) \end{array} \right.$$

Simulation of the bivariate Normal copula

Question 7

Show that this algorithm is equivalent to the Cholesky algorithm found in Question 3.

Simulation of the bivariate Normal copula

We obtain the same algorithm, because we have the following correspondence:

$$\left\{ \begin{array}{l} v_1 = \Phi(n_1) \\ v_2 = \Phi(n_2) \end{array} \right.$$

The algorithm described in Question 6 is then a special case of the Cholesky algorithm if we take $n_1 = \Phi^{-1}(v_1)$ and $n_2 = \Phi^{-1}(v_2)$. Whereas n_1 and n_2 are directly simulated in the Cholesky algorithm with a Gaussian random generator, they are simulated using the inverse transform in the conditional distribution method.

Construction of a stress scenario with the GEV distribution

Question 1

We note a_n and b_n the normalization constraints and $\mathbf{G}$ the limit distribution of the Fisher-Tippet theorem.

Construction of a stress scenario with the GEV distribution

We recall that:

$$\begin{aligned} \Pr\left\{ \frac{X_{n:n} - b_n}{a_n} \leq x \right\} &= \Pr\left\{ X_{n:n} \leq a_n x + b_n \right\} \\ &= \mathbf{F}^n \left(a_n x + b_n \right) \end{aligned}$$

and:

$$\mathbf{G}(x) = \lim_{n \to \infty} \mathbf{F}^n \left(a_n x + b_n \right)$$

Construction of a stress scenario with the GEV distribution

Question 1.a

Find the limit distribution **G** when $X \sim \mathcal{E}(\lambda)$, $a_n = \lambda^{-1}$ and $b_n = \lambda^{-1} \ln n$.

Construction of a stress scenario with the GEV distribution

We have:

$$\begin{aligned} \mathbf{F}^n(a_n x + b_n) &= \left(1 - e^{-\lambda\left(\lambda^{-1}x + \lambda^{-1}\ln n\right)}\right)^n \\ &= \left(1 - \frac{1}{n}e^{-x}\right)^n \end{aligned}$$

We deduce that:

$$\mathbf{G}(x) = \lim_{n\to\infty}\left(1 - \frac{1}{n}e^{-x}\right)^n = e^{-e^{-x}} = \mathbf{\Lambda}(x)$$

Construction of a stress scenario with the GEV distribution

Question 1.b

Same question when $X \sim \mathcal{U}_{[0,1]}$, $a_n = n^{-1}$ and $b_n = 1 - n^{-1}$.

Construction of a stress scenario with the GEV distribution

We have:

$$\begin{aligned}\mathbf{F}^{n}\left(a_{n}x+b_{n}\right) &= \left(n^{-1}x+1-n^{-1}\right)^{n} \\ &= \left(1+\frac{1}{n}\left(x-1\right)\right)^{n}\end{aligned}$$

We deduce that:

$$\mathbf{G}\left(x\right)=\lim_{n\rightarrow\infty}\left(1+\frac{1}{n}\left(x-1\right)\right)^{n}=e^{x-1}=\mathbf{\Psi}_{1}\left(x-1\right)$$

Construction of a stress scenario with the GEV distribution

Question 1.c

Same question when X is a Pareto distribution:

$$\mathbf{F}(x) = 1 - \left(\frac{\theta + x}{\theta}\right)^{-\alpha},$$

$a_n = \theta\alpha^{-1}n^{1/\alpha}$ and $b_n = \theta n^{1/\alpha} - \theta$.

Construction of a stress scenario with the GEV distribution

We have:

$$\begin{aligned}\mathbf{F}^{n}\left(a_n x + b_n\right) &= \left(1-\left(\frac{\theta}{\theta+\theta\alpha^{-1}n^{1/\alpha}x+\theta n^{1/\alpha}-\theta}\right)^{\alpha}\right)^{n} \\ &= \left(1-\left(\frac{1}{\alpha^{-1}n^{1/\alpha}x+n^{1/\alpha}}\right)^{\alpha}\right)^{n} \\ &= \left(1-\frac{1}{n}\left(1+\frac{x}{\alpha}\right)^{-\alpha}\right)^{n}\end{aligned}$$

We deduce that:

$$\mathbf{G}\left(x\right) = \lim_{n\to\infty}\left(1-\frac{1}{n}\left(1+\frac{x}{\alpha}\right)^{-\alpha}\right)^{n} = e^{-\left(1+\frac{x}{\alpha}\right)^{-\alpha}} = \mathbf{\Phi}_{\alpha}\left(1+\frac{x}{\alpha}\right)$$

Construction of a stress scenario with the GEV distribution

Question 2

We denote by **G** the GEV probability distribution:

$$\mathbf{G}(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

What is the interest of this probability distribution? Write the log-likelihood function associated to the sample $\{x_1, \ldots, x_T\}$.

Construction of a stress scenario with the GEV distribution

The GEV distribution encompasses the three EV probability distributions. This is an interesting property, because we have not to choose between the three EV distributions. We have:

$$g(x) = \frac{1}{\sigma}\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\left(\frac{1+\xi}{\xi}\right)} \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}$$

We deduce that:

$$\begin{aligned}\ell &= -\frac{n}{2}\ln\sigma^2 - \left(\frac{1+\xi}{\xi}\right)\sum_{i=1}^{n}\ln\left(1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right) - \\ &\quad \sum_{i=1}^{n}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\end{aligned}$$

Construction of a stress scenario with the GEV distribution

Question 3

Show that for $\xi \to 0$, the distribution **G** tends toward the Gumbel distribution:

$$\mathbf{\Lambda}(x) = \exp\left(-\exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right)$$

Construction of a stress scenario with the GEV distribution

We notice that:

$$\lim_{\xi \to 0} (1 + \xi x)^{-1/\xi} = e^{-x}$$

Then we obtain:

$$\begin{aligned}
\lim_{\xi \to 0} \mathbf{G}(x) &= \lim_{\xi \to 0} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \\
&= \exp \left\{ - \lim_{\xi \to 0} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \\
&= \exp \left(- \exp \left(- \left(\frac{x - \mu}{\sigma} \right) \right) \right)
\end{aligned}$$

Construction of a stress scenario with the GEV distribution

Question 4

We consider the minimum value of daily returns of a portfolio for a period of n trading days. We then estimate the GEV parameters associated to the sample of the opposite of the minimum values. We assume that ξ is equal to 1.

Construction of a stress scenario with the GEV distribution

Question 4.a

Show that we can approximate the portfolio loss (in %) associated to the return period $\mathcal{T}$ with the following expression:

$$r(\mathcal{T}) \simeq -\left(\hat{\mu} + \left(\frac{\mathcal{T}}{n} - 1\right)\hat{\sigma}\right)$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the ML estimates of GEV parameters.

Construction of a stress scenario with the GEV distribution

We have:

$$\mathbf{G}^{-1}(\alpha) = \mu - \sigma\xi^{-1}\left[1 - (-\ln\alpha)^{-\xi}\right]$$

When the parameter ξ is equal to 1, we obtain:

$$\mathbf{G}^{-1}(\alpha) = \mu - \sigma\left(1 - (-\ln\alpha)^{-1}\right)$$

By definition, we have $\mathcal{T} = (1-\alpha)^{-1} n$. The return period $\mathcal{T}$ is then associate to the confidence level $\alpha = 1 - n/\mathcal{T}$. We deduce that:

$$\begin{aligned} R(\mathcal{T}) &\approx -\mathbf{G}^{-1}(1 - n/\mathfrak{t}) \\ &= -\left(\mu - \sigma\left(1 - (-\ln(1 - n/\mathcal{T}))^{-1}\right)\right) \\ &= -\left(\mu + \left(\frac{\mathcal{T}}{n} - 1\right)\sigma\right) \end{aligned}$$

We then replace μ and σ by their ML estimates $\hat{\mu}$ and $\hat{\sigma}$.

Construction of a stress scenario with the GEV distribution

Question 4.b

We set n equal to 21 trading days. We obtain the following results for two portfolios:

Portfolio	$\hat{\mu}$	$\hat{\sigma}$	ξ
#1	1%	3%	1
#2	10%	2%	1

Calculate the stress scenario for each portfolio when the return period is equal to one year. Comment on these results.

Construction of a stress scenario with the GEV distribution

For Portfolio #1, we obtain:

$$R\left(1\mathrm{Y}\right) = -\left(1\% + \left(\frac{252}{21} - 1\right) \times 3\%\right) = -34\%$$

For Portfolio #2, the stress scenario is equal to:

$$R\left(1\mathrm{Y}\right) = -\left(10\% + \left(\frac{252}{21} - 1\right) \times 2\%\right) = -32\%$$

We conclude that Portfolio #1 is more risky than Portfolio #2 if we consider a stress scenario analysis.