

**CAS MONOGRAPH SERIES
NUMBER 15**

INTRODUCTION TO CAPITAL MODELING AND PORTFOLIO MANAGEMENT

John A. Major

Stephen J. Mildenhall

CASUALTY ACTUARIAL SOCIETY



INTRODUCTION TO CAPITAL MODELING AND PORTFOLIO MANAGEMENT

John A. Major

Stephen J. Mildenhall



Casualty Actuarial Society
4350 North Fairfax Drive, Suite 250
Arlington, VA 22203
www.casact.org
(703) 276-3100

Introduction to Capital Modeling and Portfolio Management
By John A. Major and Stephen J. Mildenhall

Copyright 2026 by the Casualty Actuarial Society

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. For information on obtaining permission for use of the material in this work, please submit a written request to the Casualty Actuarial Society.

ISBN (print edition) 978-1-7370028-8-8

ISBN (electronic edition) 978-1-7370028-9-5

Dedication

Dedicated to Donald Mango (1963–2022)

Don was a friend, colleague, and mentor to us both. His vision, guidance, and deep understanding shaped not only this monograph but our development as actuaries. A passionate educator and gifted communicator, he elevated the profession and all who knew him. We are profoundly grateful.

Contents

Monograph Editorial Board	vii
Preface	viii
About the Authors	x
1. Introduction	1
1.1. What are you getting yourself into?	1
1.2. Structure of the monograph	2
1.3. Getting started.....	3
1.3.1. The charter	3
1.3.2. Organizational factors	4
1.4. Clarifying remarks	6
1.4.1. Pricing, ratemaking, and costing	6
1.4.2. Margins and returns	6
1.4.3. Expected, plan, actual, and other flavors.....	7
1.4.4. Parts of a model	8
1.5. System overview	9
1.6. Introduction to InsCo	12
2. Business operations	15
2.1. General considerations	15
2.2. Units	16
2.3. Reinsurance	18
2.4. Asset risk and Economic Scenario Generator	19
2.5. Capital structure	20
2.6. Accounting	20
2.7. Pro formas	22
3. Capital adequacy and sources of risk	24
3.1. Capital adequacy	24
3.2. Value at Risk and Tail Value at Risk	27
3.3. Risk analysis	29
3.4. Algorithms.....	31
3.4.1. Algorithm to evaluate expected loss for discrete random variables ..	31

3.4.2.	Algorithm to evaluate VaR for a discrete distribution	33
3.4.3.	Algorithm to evaluate TVaR for a discrete distribution	33
4.	Pricing and allocation	34
4.1.	Pricing the whole	35
4.2.	Allocation to units: Overview	39
4.3.	Four technical adjustments	40
4.4.	Pricing layers in total	42
4.5.	Pricing units via pricing layers	48
4.6.	Five representative distortion functions.....	50
4.7.	The industry standard approach and its problems	52
4.8.	Pricing events: An explanation	56
4.9.	Summary	57
4.10.	Algorithms	58
4.10.1.	Algorithm to evaluate an SRM on a discrete random variable	58
4.10.2.	Algorithm to compute the linear NA for discrete random variables	59
5.	Applications of pricing and allocation	61
5.1.	Performance assessment.....	61
5.2.	Reinsurance decision-making	62
5.3.	New business pricing	64
5.4.	Mergers and acquisitions.....	66
5.5.	Portfolio optimization	68
5.5.1.	Theory.....	68
5.5.2.	InsCo optimization.....	71
5.5.3.	Summary	75
6.	Selecting and calibrating an SRM	76
6.1.	The envelope of possibilities and the five representatives	76
6.2.	Required premium by distortion	78
6.3.	Zeroing in on a distortion function	79
6.4.	Embracing the range	80
7.	Evaluating models.....	81
7.1.	Evaluating your model	81
7.2.	Evaluating third-party models	82
7.3.	Evaluating modeling platforms	83
8.	Advanced topics.....	85
8.1.	Dealing with uncertainty	85
8.2.	The Iman-Conover method	86
8.2.1.	Loss or loss ratio?	87
8.2.2.	Example 1: A simple example	87

8.2.3.	Different reference distributions	90
8.2.4.	Copulas as reference distributions	91
8.2.5.	Iman-Conover compared with the normal copula method	91
8.2.6.	Block Iman-Conover (BIC)	93
8.2.7.	Example 2: Multistate, multiline correlation	94
8.2.8.	Theoretical derivation of the Iman-Conover method	97
8.3.	Allocating capital, if you must	99
8.4.	Comparisons with Bodoff's layer capital allocation	105
9.	Calculations with aggregate	107
9.1.	aggregate and Python	107
9.2.	aggregate and R	107
9.3.	Reproducing the code examples	107
9.4.	Reproducing the discrete example	108
9.4.1.	Input data and setup	108
9.4.2.	Premiums by distortion	112
9.4.3.	Reinsurance analysis	113
9.5.	A more realistic example	114
	Appendices	121
A.1.	Table of symbols	121
A.2.	Actuarial Standards of Practice	123
A.3.	Glossary	124
	References	127

Monograph Editorial Board

Brandon Smith, Editor in Chief

Emmanuel Bardis

Marco Dattilo

Marco De Virgilis

Scott Gibson

Michael Henk

Ali Ishaq

Erin Lachen

Joseph Lindner

Wangsun Xia

Janice Young

Yi Zhang

Yuhan Zhao

Yuanshen Zhu

Preface

Writing a monograph on capital modeling is a humbling experience. Capital Modeling, also known as Dynamic Financial Analysis (DFA), Enterprise Risk Analysis, or Internal Risk Modeling, is at least 120 years old. It began in 1903, when Swedish actuary Filip Lundberg introduced the stochastic compound-Poisson risk model to describe how an insurer's surplus evolves over time. The literature is vast, thousands of pages across hundreds of books and articles. How can one condense that all down to a comprehensive how-to manual? The answer is, one cannot, and we do not try. Instead, we attempt to sketch out what needs to happen, and provide sufficient references to the existing literature so that the reader can obtain whatever detail is needed.

Where we have introduced some innovation is at the very end of the process—capital allocation, or, more properly, the allocation of cost of capital. Here, we explore in detail the (unfortunate) consequences of assuming a fixed cost of capital rate and offer a TVaR-based generalization in its place: Spectral Risk Measures, also known as Distortion Risk Measures, and their Natural Allocation. Yet we cannot claim originality here, either, as we are leaning heavily on another decades-long vein of actuarial research.

With allocation, we deliberately adopt a single-period perspective. We do not address the time required for claims to develop to ultimate, nor do we treat the insurer as a going concern. The first omission aligns with most regulatory capital models, which focus on one-year changes in reserves and new accident year loss bookings. They account for reserve risk, but not the full run-off risk. The second reflects our desire to avoid getting bogged down in what David Babbel (discussion in Reitano (1997)) calls the “quagmire of equity valuation”—a topic one of us has written about already and the other chooses to avoid.

It is perhaps a fitting irony that the quantification of risk is riddled with uncertainty. Capital modeling is part science, part dark art—and its real value often isn't the final number, but the insight gained along the way. While some may chase the best estimate, models are most useful when they show a range of outcomes and the assumptions behind them. Some inputs are squishy, like risk appetite, and it's often easier to rule out wrong answers than pin down the right one. Think of capital modeling less as a calcu-

lator and more as a compass: it won't give you the answer, but it helps you head in the right direction.

We draw heavily from our book Mildenhall and Major (2022). You do not need the book to implement the mechanics presented in this monograph; we wrote it to stand alone. However, it is a good source for the whys and wherefores, and advanced topics.

Finally, capital modeling is a participation sport not a spectator sport. We recommend strongly that you open your favorite spreadsheet or software tool and reproduce the simple examples. Generally, as the models become more realistic, they scale up in terms of number of rows and columns but not in terms of conceptual complexity. Both authors have used the basic methods outlined on 10,000+ simulated events with 100+ units and obtained satisfactory results—though this should not be attempted in a spreadsheet unless you are exceedingly patient. We apologize that straightforward spreadsheet concepts, such as the sum product of two columns or a weighted average, become intimidating mathematical formulas, but since we want to be explicit, we need to present the formulas and there's no sidestepping this. The reader will find it helpful to see through the formulas to their simple spreadsheet form, best done hands-on. Lastly, this paragraph should not be taken as an endorsement of spreadsheets. You should learn to program!

About the Authors

John A. Major has served as a research leader and data scientist in diverse insurance contexts, contributing to the state of the art in areas such as claim fraud detection, insurance-linked securities, terrorism risk, and catastrophe modeling. His publications in over a dozen books and journals have been cited in hundreds of scholarly articles. He is an Associate of the Society of Actuaries and holds a Master's degree in mathematics from Harvard University.

Stephen J. Mildenhall has extensive general insurance experience, having worked in primary and reinsurance pricing, broking, and education since 1992. He is a Fellow of the Casualty Actuarial Society, an Associate of the Society of Actuaries, and holds a PhD degree in mathematics from the University of Chicago.

1. Introduction

1.1. What are you getting yourself into?

Congratulations! You have been tasked with developing a capital modeling system! But what does this mean? Maybe you were lucky enough to be handed a detailed specification of what this capital model is supposed to do and who is going to make use of its outputs and how those outputs will affect business operations. Or perhaps not.

But what is a capital modeling system? What is it for? And how do you build one?

Actuarial standards define **capital** as “the funds...over and above those funds backing the liabilities,” where “funds backing the liabilities” are amounts booked to support incurred liabilities (loss reserves) or future liabilities (unearned premiums) (Actuarial Standards Board 2011). Capital is then a cushion in excess of those amounts. This definition highlights that capital is an accounting concept whose value will vary with the accounting standard. Since the distinction between capital and reserves is blurry, it is often more productive to focus on all available assets. Pricing then splits assets between policyholder-provided premium and reserves and owner-provided capital.

The modeling standard (Actuarial Standards Board 2019) defines a **model** as “a simplified representation of relationships among real-world variables, entities, or events using statistical, financial, economic, mathematical, non-quantitative, or scientific concepts and equations.” A key word here is *simplified*; as Gabaix and Laibson (2008) point out, “a perfect replica of the earth would reproduce every contour, but such a representation would be impractical.” Capital modelers should take to heart the admonition attributed to Einstein to make things “as simple as possible, but no simpler.” Any increment of complexity should materially improve the utility of your model. Modelers face many pressures to build granular, complex models but receive no countervailing rewards for accuracy and practicality.

Today, computers are ubiquitous in modeling, and this monograph will focus exclusively on quantitative models embedded in computer systems.

A capital modeling system may be expected to support:

- risk identification and analysis
- product pricing (ratemaking)

- rating and underwriting
- reinsurance decisions
- business unit performance assessment
- portfolio optimization
- valuation of a potential spinoff or acquisition
- regulatory compliance
- rating agency discussions
- strategic planning

We will assume in this monograph that your goal is, ultimately, a comprehensive system that can support all of the above activities. Of course, as you build it, you will need to prioritize interim goals and use cases, but directionally this is where we assume you are heading. For modelers in the US, questions of pricing are more important than questions of capital adequacy, as the latter constraints tend to be externally imposed. In the EU, however, an internal model may be called on to determine capital adequacy. There, regulators issue extensive guidance on model building that obviously trump any conflicting advice we offer.

The system's core modeling function is to *mirror and tie into the plan*. It is first a system to support and enhance the existing planning process and second to *quantify risk around the plan*. Support for the other listed tasks stems from this capability to represent risk.

1.2. Structure of the monograph

The remainder of the Introduction helps you with **Getting started**, addressing high-level goals and operational issues to get you off the ground in Section 1.3, making some **Clarifying remarks** in Section 1.4, providing a **System overview** in Section 1.5 that outlines a generic capital modeling system and that forms the basis for subsequent discussions of functionality, and in Section 1.6, **Introduction to InsCo**, presenting a highly simplified example that is used throughout for illustration.

The rest of the monograph is structured as follows:

- Chapter 1 **Introduction** is this chapter.
- Chapter 2 **Business operations** discusses the largest module in the capital modeling system. This is where the financial fortunes of the firm are simulated. Building or acquiring and implementing this module will absorb most of your resources. This function is also where decades of training, experience, and literature can be brought to bear. In this chapter we will provide further pointers into that literature.
- Chapter 3 **Capital adequacy and sources of risk** discusses the second module, which is smaller and easier to implement. It depends on the outputs from Business Oper-

ations and reflects well-understood principles of risk measurement. Basic risk measurement as well as regulatory aspects are discussed in this chapter. We also discuss risk analysis applications. The simple example continues. At the end of the chapter, algorithms for computing expected value, quantiles (VaR), and TVaR on discrete distributions are presented.

- Chapter 4 **Pricing and allocation** is where we introduce new ways of thinking, specifically Spectral Risk Measures (SRMs) and the Natural Allocation thereof. In this chapter, InsCo is treated with a complete pricing and allocation exercise. Along the way, we discuss the industry standard approach to capital allocation and why it is problematic. At the end of the chapter, algorithms for computing SRMs and the Natural Allocation on discrete distributions are presented.
- Chapter 5 **Applications of pricing and allocation** use the outputs of the Capital Adequacy and Pricing and Allocation modules in performance assessment, new business pricing, reinsurance decision-making, portfolio optimization, and mergers and acquisitions.
- Chapter 6 **Selecting and calibrating an SRM** discusses how choice of an SRM impacts the various portfolio components and presents an approach to finding an acceptable solution. It also discusses how to leverage the range of consistent outcomes and understand their implications on risk appetite.
- Chapter 7 **Evaluating models** discusses how to evaluate your own model, third-party models, and modeling platforms.
- Chapter 8 **Advanced topics** deals with uncertainty and allocating capital (if you must) and a clever technique to implement correlation.
- Chapter 9 **Calculations with aggregate** shows how to use the open-source `aggregate` package to work the examples in this monograph.
- The **Appendices** include a table of symbols, list of relevant Actuarial Standards of Practice, and a glossary.

1.3. Getting started

1.3.1. The charter

The maxim “If you build it, they will use it” does not apply to capital models. Successful models are built in response to business needs, with high-level direction set by a senior business leader. A **charter** can be a valuable tool in crystallizing that direction. You should have some more-or-less formal document answering design questions such as:

- Why are we building a capital modeling system?
 - Is this driven by regulatory concerns, dissatisfaction with the current planning process, or something else?

- What types of decisions will it support?
 - See examples in Section 1.1.
- Who is going to use this model?
 - Who is going to run it, hands on?
 - How often will it run (daily, weekly, monthly, etc.)?
 - Who is going to be in a position to set inputs or pose queries?
 - Who is going to receive reports?
 - What is the users' level of sophistication, in terms of what they are likely to ask and what sort of output or report they are willing and able to absorb?
 - Is the model intended for senior staff to be reviewing outputs or junior staff to be integrating outputs with their own workflow? This may differ across functional units.
- When will the outputs be used?
 - During the planning process, presumably, but probably other times as well.
- Whose business processes will be affected by the system?
- How will outputs be integrated into planning and underwriting?
- How will we know if we are successful?
 - Define metrics such as use and impact.
 - What stakeholders need to be in agreement to make the system successful?

Ideally, this document will bear a cover memo endorsed by some sufficiently high-ranking executive. At the very least, it must be widely exposed to senior management and key stakeholders (typically Lines of Business) via road shows.

How will you know if you are successful? Mainly, by seeing your model get used, i.e., “fully embedded into the risk strategy and operational processes of the insurer” (IAIS 2008). This means, among other things:

1. Management has an overall understanding of what your model does, in terms of input versus output.
2. The model is used in the planning cycle; results are regularly discussed and reviewed.
3. Decisions that affect the firm's risk profile usually refer to model results.
4. Decision-makers ask when the next set of outputs will be ready.

The European Insurance and Occupational Pension Authority (EIOPA (2009), title I, chapter VI, section 4, article 120; EIOPA (2015) title I, chapter 6) elaborates on these points in the context of the use test.

1.3.2. Organizational factors

Your project plan needs to address organizational issues. What resources will be available for initial development, ongoing maintenance, and operations? What actuarial and pro-

gramming staff will need to be involved? What data are required? What funds are needed to purchase software? What hardware is needed?

Don Mango addresses organizational and human factors related to implementing an internal risk model (IRM) in Brehm et al. (2007, Section 3.1). He covers the following areas:

1. **Startup staffing:** This step includes considering the organizational chart, business functions represented, resource commitments, roles, and responsibilities. He recommends that system implementation is “the establishment of a new competency” and calls for the transfer of full-time employees or even adding new hires. Control of inputs and outputs of the model needs to be taken as seriously as general ledger or reserving systems.
2. **Purpose and scope:** He suggests initially to cover only the prospective underwriting year and to quantify variation around that plan; later, expand the scope to include reserves, assets, and operational risks with more formal probability distributions.
3. **Parameter development:** Data quality issues will require expert judgment from underwriting, claims, planning, and actuarial staff. “Develop a systematic way to capture expert risk opinion.” Correlation is as important as it is problematic.
4. **Software:** Evaluate available software solutions, and decide whether to build or buy. This decision depends on the strength and skills of the modeling team.
5. **Validation testing:** This will occur over time and requires significant planning. Reach out to affected parties and make sure they understand what validation does and how it works.
6. **Rollout:** Early in the process, secure top management backing for prioritization, communications and training, and pilot testing.
7. **Integration and maintenance:** Of course, he recommends mandatory inclusion of model reports as part of decision support in the planning cycle. Make major updates no more frequently than semiannually. Maintain centralized control of inputs and outputs.

Emma et al. (2000, Chapter 6) discuss dangers and pitfalls. We can recast that advice as things to do:

1. Make sure to get sufficient senior management support.
 - Get a charter!
 - Make sure the effort is going to support a real business use.
2. Attend to communication between the team and key stakeholders.
 - The model design may be quite technical. Make sure stakeholders understand enough about the model to be informed on what it can and cannot do, as well as what is required to develop and maintain it.

- Use the same language as stakeholders. This is especially important when talking to professionals like accountants.
 - Allocate enough time to communication.
3. Make sure model results can be integrated into planning.
- Results must reconcile to plan.
 - Results must emerge in time to be used in the planning process.

Planning the development of a capital modeling system is very much like planning any other significant software system. Treat it with the same level of care.

1.4. Clarifying remarks

1.4.1. Pricing, ratemaking, and costing

Throughout the monograph we talk about *pricing*, using phrases like the “required price” to mean a model indicated price. For actuaries working outside highly regulated lines, we believe our terminology is standard and immediately comprehensible. However, actuaries working in regulated lines may be sensitive to the distinction between costing, ratemaking, and pricing. Cost estimates consider only loss-related differences between insureds. Ratemaking allows behavioral differences to creep into premiums, considering price elasticity and propensity to shop in determining a profit maximizing rate, i.e., allowing price optimization. Finally, price is usually reserved for a market determined sold rate. Acceptable prices may include corporate strategic objectives aside from considerations of the insured’s risk. We do not explicitly consider price optimization or other strategic pricing objectives and are therefore really engaged in costing, but we find the word “pricing” preferable and trust that no confusion will arise.

1.4.2. Margins and returns

Much actuarial ink has been spilled on the distinction between *margins* and *returns* (McClenahan 1999). The situation is straightforward. Insureds care about premiums and hence margins to premium. Investors care about capital and hence the return on capital. Insurers and regulators care about both. Capital is a cost for an insurer, hence they speak of the *cost* of capital. Investors purchase the securities (equity, bonds) insurers use to finance capital, and they are interested in the *return* on their investment. Cost of capital is almost always *ex ante* and expected; return on investment can be expected or *ex post* and actual. The rate regulator focuses on premium margins and the embedded cost of capital, whereas the solvency regulator monitors actual returns and the risks taken to achieve them.

Premium leverage forms the bridge between premium and capital. The comparison in Table 1.1 helps illustrate this bridge. It contrasts the typical economics of an attritional,

Table 1.1. Comparison of loss ratio, leverage and return for attritional and catastrophe type units.

Unit Type	Measure	Value	Qualitative	Quantitative
Attritional	Loss ratio (%)	High	Cheap insurance	90s
	Leverage	High	Efficient pooling	2:1
	Return (%)	High	High returns	Teens to 20s
Catastrophe	Loss ratio (%)	Low	Expensive insurance	30s
	Leverage	Low	Inefficient pooling	1:5
	Return (%)	Low	Low returns	Single digit

low-risk unit with those of a high-risk, catastrophe-exposed one. The value column is particularly important: it clarifies a common point of confusion. Attritional insurance is typically inexpensive in the sense it is written at a *high* loss ratio. It is also written at *high* leverage, often producing a *high* equity-like return on capital. Conversely, catastrophe risk is expensive and written at a *low* loss ratio. It is difficult to pool efficiently and as a consequence is written at *low* leverage, resulting in *low* debt-like returns, as seen in the catastrophe bond market. This can seem paradoxical: business that is *more* risky from an insurance perspective yields a *lower* return on capital, and vice versa. The paradox is resolved by recognizing the difference between risk to the *insurer* and risk to the *investor*. Capital supporting attritional risk faces a roughly 50% chance of short-term loss, whereas capital supporting pure catastrophe risk has a very low probability of loss. The loss-given-default characteristics also differ substantially. Last, it might seem we should use margin in this table to measure expense rather than loss ratio. We don't because margin is negatively correlated with loss ratio, so using margin would break the clean high/high/high vs. low/low/low symmetry of the table.

1.4.3. Expected, plan, actual, and other flavors

When we talk about pricing, terms like “loss” and “margin” usually mean *expected* loss or margin. It is helpful to distinguish several different pricing viewpoints (Robbin 1992). A premium or rate can be:

1. **Filed:** Incorporated into filed rates in a regulated line.
2. **Required or Target:** Premium required to hit a (corporate) pricing target.
3. **Plan:** Reflects adjusted targets actually embedded in a plan.
4. **Indicated:** Produced by an actuarial pricing model or process.
5. **Priced:** Produced by an actuarial underwriting process.
6. **Quoted:** Actually released into the market as a bindable rate.
7. **Sold:** The premium actually sold in a transaction.
8. **Market:** An estimate of sold rates.

The same terms can be applied to margins. These terms are almost always meant prospectively and are *ex ante*. In addition, for margins there is the important:

9. **Actual:** Achieved margin reflecting developed ultimate loss experience.

Actual is obviously after the fact or *ex post*.

Our pricing discussions primarily focus on setting some flavor of required premium in the context of a corporate or business unit plan. We often use the term “required” as a synonym for “target” or “indicated,” i.e., a price computed by a model in order to meet some specified objective.

1.4.4. Parts of a model

Before describing what should be in a model, some distinctions should be kept in mind. There are various *levels* to modeling, starting from the computer infrastructure and extending all the way up to the application being run by the end-user. Table 1.2 lays out these levels and associated concepts.

Table 1.2. Model components from hardware to output.

Concept	Examples
Levels include	
Computer hardware	Mac, PC, server, workstation, cloud (AWS, Databricks)
Operating system and system utilities	macOS, Windows, Linux
Programming language	Python, R, Java, Visual BASIC, C++
Modeling platform	Excel, ReMetrica, IGLOO, MetaRisk, Oasis, your own custom “empty” spreadsheet or function library
Model	Spreadsheet Rating agency capital models, Solvency II standard calculation Specific peril cat model, e.g., vendor US Hurricane model
Models include	
Modeling techniques	Simulation, calculation
Structural parameters	Protected cells, e.g., unit loss volatility and correlations Cat model damage curve
Business parameters	User input cells, e.g., unit plan premium and loss ratios Cat model property portfolio
Raw outputs	Cat model event or yearly loss tables, event timeline database
Output exhibits	Probable maximal loss table, summary statistics, pro forma financials

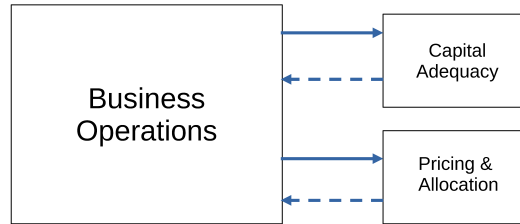


Figure 1.1. Three modules of a capital modeling system.

Structural parameters are changed infrequently. Business parameters are changed frequently, especially during planning when it is standard to run many different iterations.

1.5. System overview

There is considerable variation in how a capital modeling system could be designed and implemented. In this section we provide a conceptual template to follow.

We can mentally decompose a capital modeling system into three components shown in Figure 1.1. By far, the biggest and most complex module represents the stochastic business of an insurer. This is labeled **Business Operations** in the diagram. A capital modeling system must represent some or all of the firm's portfolio of insurance liabilities.

There is a wide range of sophistication possible to this representation, depending on the particulars of the purpose for which the system is to serve. The core concept is the loss that the portfolio might incur. This loss, being *a priori* uncertain, is a random variable. It might be represented by a distribution of losses or by a set of loss scenarios, which we will call events. The output could include just the losses, or it could include a complete set of financial statements—balance sheet, income statement, and cash flows—all of which ultimately depend on the losses. If the latter, there are numerous accounting standards that could be followed. These possibilities are discussed in Chapter 2.

Output from the Business Operations module feeds into the much simpler **Capital Adequacy** and **Pricing & Allocation** modules. This is signified by the solid arrows. Dotted arrows signify that these two modules, in turn, affect business operations in the real world and hence, indirectly, the Business Operations module. Indeed, the real-world effect of these modules is the *raison d'être* of the entire capital modeling system.

Figure 1.2 expands the Business Operations module.

Starting at the left of the diagram is the **Units** module, with inputs labeled **U**. We use the term **unit** to refer to a component of the portfolio. A unit can be a single policy or a group of policies defined by line of business, geography, branch, business unit, or other characteristics. A unit can also represent the segmentation between ceded and assumed reinsurance.

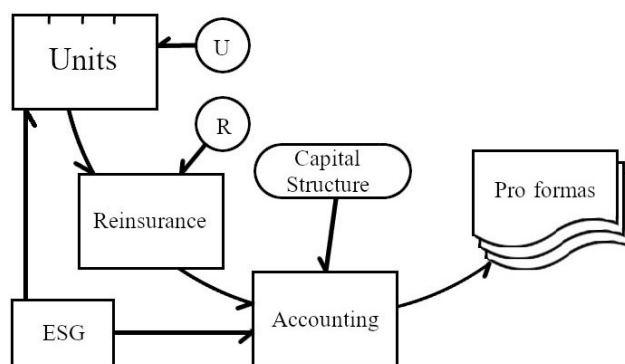


Figure 1.2. Internal structure of Business Operations module.

The output of Units is a representation of *experience*, which means paid loss, at the very least, but could also include a time series of loss developments, reserves, etc. Input parameters are things like exposure and possibly shape parameters for loss distributions.

These loss outputs are likely to be correlated (not independent) across the units either by common economic or social factors or enforced correlation via copulas, Iman-Conover sampling (Section 8.2), etc.

If investment experience is included as one (or more than one) unit, then input from the **Economic Scenario Generator** (labeled ESG) will be important. This is an optional module that represents external economic factors such as interest rates, returns on various asset classes, rates of monetary inflation, unemployment, etc.

The **Reinsurance** module, with inputs labeled **R**, modifies the gross experience of the units. This includes all outward (ceded) reinsurance. Any inward (assumed) reinsurance would be represented among the units. The output, net experience, includes the effect of any retrocession.

Accounting translates experience realizations into **Pro formas**, i.e., financial statements. This output from the Accounting module is a core element of the system. Multiple occurrences of pro formas, generated on the fly or possibly stored in a database, are used in stochastic simulation, whereas single occurrences are used for analyzing specific events. The level of detail here must match the specification of the units.

The system is likely to include the input labeled **Capital Structure**. This specifies, at the very least, the total amount of assets available to pay claims. It may also specify details of noninsurance liabilities, such as bonds, preferred stock, and equity issued by the firm. The Economic Scenario Generator may have some influence on the outward cash flows—which reduce assets—from preferred and common stock dividends. The Accounting module would take into account the fact that losses in excess of total assets lead to *insolvency* and haircuts in claim payments. If there is no capital structure information at all, asset limitations are ignored and assets are in effect assumed to be infinite.



The **Capital Adequacy** module, with inputs labeled **C**, operates on the pro formas and determines the minimal amount of assets to satisfy requirements from regulators, rating agencies, and the risk committee of the firm's board of directors. Such determinations could include a simple Value at Risk (**VaR**) calculation or more complex Best's Capital Adequacy Ratio (**BCAR**) calculation. The module also includes risk reports—analysis of what is driving the need for assets. The input parameters can be regarded as the firm's **risk tolerance**, specifying what level of risk is acceptable for a given amount of capital or vice versa. Outputs feed back into the real world and implicitly affect the capital structure, hence the dotted line. This is discussed in more detail in Chapter 3.

Allocation takes the required portfolio premium as given and allocates the premium (equivalently, margin, loss ratio) to the individual units. This information, presumably, has real-world impacts on the units via performance management, risk management, underwriting actions, etc. Impact is represented by the dotted arrows to Units.

1.6. Introduction to InsCo

InsCo is our one-period simplified model of insurance company operations.

We will use an extremely simple toy example to illustrate the operation of the system presented in Section 1.5. It deliberately abstracts away from much of the complexity, detail, and nuance of the real world. Readers should be able to reproduce the example easily in a spreadsheet.

A more complex and realistic example can be found in Section 9.5, and more in Mildenhall and Major (2022). Those, as well as this simple example and even more complex portfolios, can all be analyzed with `aggregate`, an open-source Python library (Mildenhall 2024) that “builds approximations to compound (aggregate) probability distributions quickly and accurately.” Because `aggregate` uses Fast Fourier Transforms, it “delivers the speed and accuracy of parametric distributions to situations that usually require simulation” and is therefore much faster than simulation. All the analytical techniques discussed in this monograph can be implemented in `aggregate`.

InsCo is a limited liability company that intermediates between insureds and investors. InsCo’s customers are **insureds** (policyholders) who are subject to risks they wish to insure. Insureds who use insurance for risk transfer or financing are sensitive to insurer quality and possible default because it correlates with their own misfortune.

Insurance legal entities serve two principal purposes. First, they provide statutory insurance such as mandatory automobile liability. Second, they allow insureds to pool together and benefit from diversification without requiring onerous bilateral contracts. They do this through insolvency rules, which provide the framework under which unrelated insureds interact in the unlikely event of an insolvency.

InsCo comes into existence at time $t = 0$ and lasts for one period. InsCo has no initial liabilities. At $t = 0$, it writes one or more single-period insurance contracts and collects premiums from its insureds.

When InsCo writes a policy, it collects premium at $t = 0$ and earns it over the period. We assume all other transactions occur at the end of the period. Therefore all the premium is earned and available to pay claims at $t = 1$. If InsCo’s ending assets a are insufficient to pay the claims, then it defaults.

InsCo has promised to pay policyholders claims under various contingencies, with the aggregate promise represented by the random variable $X \geq 0$. If $X > a$, then only a gets paid out, i.e., the actual payments are the minimum of X and a , which we write as $X \wedge a$. We assume the probability distribution of X is known.

InsCo is owned by **investors** who provide risk-bearing capital. Investors are also risk averse. The market structure is displayed in Figure 1.4.

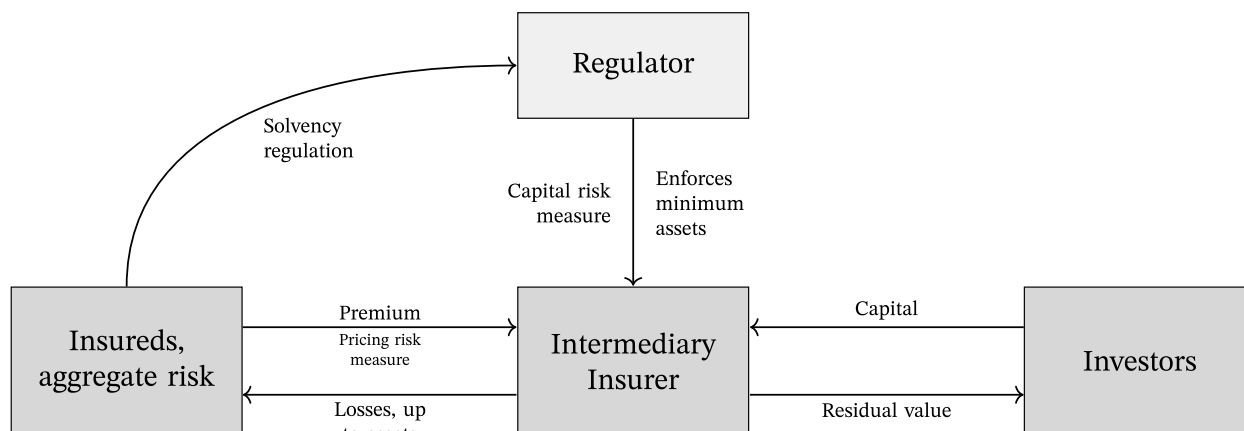


Figure 1.4. InsCo, insureds, regulators, and investors: roles and cash flows.

At time $t = 0$, while collecting premiums, InsCo also raises capital from investors by selling them its uncertain $t = 1$ residual value. That is, at time $t = 1$, InsCo pays any claims due in the amount of $X \wedge a$ and pays any residual value $(X - a)^+$, if it exists, to its investors as return of capital plus a dividend or investment return. If InsCo's ending assets are insufficient to pay the claims, $X > a$, then it defaults. Investors have limited liability: they may lose their original investment but owe nothing more.

Premiums cover expected losses and loss adjustment expenses, as well as the cost of capital including frictional capital costs. All other expenses are outside our model.

Symbolically, at time $t = 0$, InsCo collects premiums P from policyholders and capital Q from investors. These are the only sources of funds and comprise the total assets via the **funding equation**:

$$a = P + Q. \quad (1.1)$$

Two important questions arise from InsCo's promises to pay.

1. Are there sufficient assets to honor those promises?
2. Are investors being adequately compensated for taking on those risks?

Crucially, we need to talk about not one but *two* different risk measures to answer these questions.

Question 1 concerns risk tolerance and is answered by the Capital Adequacy module. It determines the assets necessary to back an existing or hypothetical portfolio at a given level of risk. This exercise can also be reverse engineered: given existing or hypothetical assets, what constraints on business does the risk tolerance entail? Alternatively, given business and capital, what is the implied risk tolerance?

Assets a and liabilities X are related by some rule driven by a combination of regulatory authorities, rating agencies, and InsCo's own internal risk management policies, representing a risk tolerance. Such a rule we call a **capital risk measure**, and we may write a as a functional $a(X)$. Value at Risk (VaR) or Tail Value at Risk (TVaR) at some high confidence level, such as 99.5% or 1 in 200 years, are both popular, but other possible measures exist; see Chapter 3. As a first approximation, we may take it that a is sufficient to avoid insolvency altogether, i.e., in all events, all claims are paid.

Question 2, answered by the Pricing module, concerns how that asset amount a is to be split between premium P and capital Q (Equation 1.1); this is quite different from determining a . It is about risk pricing or risk appetite. We must determine the expected margin insureds need to pay in total to make it worthwhile for investors to bear the portfolio's risk. Such a rule we call a **pricing risk measure**, and we may write premium as a functional $P = \rho(X)$.

2. Business operations

Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.

—George E. P. Box

2.1. General considerations

Modeling the insurance business operations—in terms of cash flows and other accounting concepts—is the most significant part of capital modeling work. It would be impractical to try to explain this in detail here; multiple book volumes would be required. Indeed, much of the actuarial syllabus is devoted to knowledge that could be applied to this effort. Instead, this chapter offers an overview using a simple example and adds pointers into the existing literature.

After you have addressed the broad questions posed in Section 1.3, you will have to drill down into more specific design issues such as:

- Will the model be deterministic or stochastic?
- What is the planning time horizon: one year versus an ongoing concern?
- What is the calculation time horizon: one year versus to ultimate?
- What is the time granularity: annual, quarterly, monthly, etc.?
- Will it deal with the runoff of current business only, or will it model an ongoing concern?
- How will it treat reserve risk?
- How will it treat asset risk?
- How will it treat investment income (risk-free rate, crediting, capital charge, simplifications from discounting)?
- How will multiple related legal entities be reflected in the model?
- What is the balance of variables between user-selected inputs (e.g., plan premium) and modeler-selected inputs (e.g., correlation or severity)?
- What accounting and regulatory frameworks and constraints are supported (US Statutory, Solvency II, US GAAP, IFRS 17, rating agencies)?
- How will the model relate to other corporate systems?

See Emma et al. (2000) Chapter 4 for some more considerations on these dimensions of functionality; Chapter 7 goes into more detail about outputs and reports. Chapter 5 presents a classification of risks:

- Asset risk (default, market value, liquidity)
- Obligation risk (reserve, premium, loss projection, catastrophe, reinsurance, and expense)
- Interest rate risk (asset, obligation, and net cash flows)
- Mismanagement risk

How deeply each is to be modeled is part of the requirements to be determined. Emma et al. (2000) discuss the first three at length, but not the last.

Subsequent sections in this chapter elaborate on some of these issues.

2.2. Units

Business operations start with the individual business units. These are the sources of premiums, losses, and some of the expenses. Expect to devote considerable effort to modeling losses; your loss model implementations will live in the Units module.

The design issues at this stage include:

- Granularity of represented components, i.e., breakdowns by geography, line of business, etc.
- Granularity of frequency and severity assumptions.
- Inclusion of parameter uncertainty, see Section 8.1.
- Interactions among variables.
- Feedback loops, i.e., embedded automatic conditional decisions. For example, future rate changes might be calculated based on emerging loss ratios.

The existence of mismanagement risk (Section 2.1) suggests that building in feedback loops—“management in a box”—might be too optimistic. We strongly recommend against trying to do this. Adjusting rates to follow inflation is reasonable because this is very likely to happen in reality. Adjusting rates on the basis of a previous random good or bad year is not reasonable because it is highly speculative as to what will happen in reality. Instead, have the model in effect state “if nothing is changed, this is what would likely happen next....”

Daykin et al. (1993) include the following chapters relevant to the Units submodule:

- 2 - Frequency
- 3 - Severity
- 4 - Compound distributions
- 5 - Simulation

- 9 - Extended time horizons, including reserves
- 10 - Premiums and markets
- 11 - Expenses
- 12 - Cycles and growth

Brehm et al. (2007) include:

- 5.1 - Frequency and severity distributions
- 5.2, 5.3 - Reserves
- 5.5 - Underwriting cycles

Nearly the entirety of Klugman et al. (2019) is devoted to selecting, fitting, and applying mathematical models of losses.

If you are modeling reserves, you must accurately represent your own firm's methodology. The literature on setting reserves is enormous, with Friedland (2010) being a good place to start. Seventeen text references on advanced reserving are listed at Casualty Actuarial Society (2022). Wüthrich and Merz (2015) is a rich source of information on stochastic reserve models. Szkoda et al. (1995) has a list of about 200 reserving considerations. Actuarial Standards Board (2007) is also relevant.

Modeling the correlation between units' loss experience can be particularly tricky. Reshuffling simulated loss outcomes to obtain a desired correlation matrix can be done by the Iman-Conover methods; see Section 8.2.

Aas et al. (2009), Embrechts and Puccetti (2010), and Brehm et al. (2007, Section 3.3) discuss copulas, which are mechanisms for introducing dependency that goes beyond simple correlation. For example, the routine loss bodies of two loss distributions can be uncorrelated while the high-loss tails become correlated.

Another technique for inducing dependence is the use of common causes. If there are underlying phenomena that influence losses in different units, then modeling those first, and the conditional distribution of unit losses second, can induce the desired dependence. This is most commonly seen in catastrophe modeling but can also appear in a technique called timeline simulation (Brehm et al. 2007, Section 3.4), where specific events are modeled as occurring at specific points of time.

We now illustrate the workings of the Business Operations module with our InsCo example.

In a real engagement, we would start with historical data on exposures, premiums, and losses. They would be on-leveld to represent the current (or anticipated next accounting year) situation. Probabilistic models would be fit, and a simulated sample of future possible outcomes generated. This work is covered early in the actuarial exam syllabus, and we imagine it has already been done.

InsCo has units $i = 1, \dots, m = 3$, called A, B, and C. We represent their random experience X^i by a table consisting of 10 outcomes with equal probabilities $p_j = 1/10$, $j = 1, \dots, n = 10$, with components X_j^i and portfolio totals X_j . The exact assumptions are detailed in Table 2.1.

Table 2.1. InsCo loss probability mass function by unit with expected (EX) loss, CV loss, and plan premium.

Event j	Unit A X_j^1	Unit B X_j^2	Unit C X_j^3	Total X_j
1	15	7	0	22
2	15	13	0	28
3	5	20	11	36
4	7	33	0	40
5	13	20	7	40
6	5	27	8	40
7	15	16	9	40
8	26	19	10	55
9	17	8	40	65
10	16	20	64	100
EX	13.4	18.3	14.9	46.6
CV	0.453	0.412	1.324	0.455
Plan premium	13.9	18.7	19.6	52.2

There is some correlation among the simulated experience of the units. This correlation might have been imposed by a sampling technique involving copulas, but here we simply present the results. Table 2.1 represents the loss portion of the Unit module. Statisticians would call this table a discrete joint distribution, sample, or empirical distribution. Cat modelers might call it a **probabilistic database**. It shows expected loss for the portfolio is 46.6, plan premium is 52.2, and hence plan profit, or margin, is 5.6.

2.3. Reinsurance

Assumed reinsurance, i.e., selling reinsurance or retrocessional cover to other insurance companies, should be treated as another unit, like homeowners or auto. Ceded reinsurance can be treated in one of two ways.

The simplest way is to treat reinsurance as a negative unit in parallel with the underlying unit. The premiums it receives are booked as negative (because they are really paid), and the losses it pays are booked as negative (because they are really received). For example, say there is the homeowners line of business being protected with catastrophe reinsurance. Then, the homeowners gross premiums and losses are posted to one unit and the reinsurance ceded premiums and ceded losses are posted as negatives to a parallel

ceded homeowners unit. The algebraic sum of premiums (respectively, losses) constitute the net homeowners position, which is not explicitly represented.

The more complicated way, and that implied by the design presented in Section 1.5, is to have the operation of reinsurance modeled downstream of the units. This is recommended because there are many situations, such as corporate catastrophe covers, that impact multiple units.

In the end, it is simply a computational design choice about where calculations occur. Mathematically, the same gross-to-net transformation is being modeled either way.

The capital modeler may encounter two different types of reinsurance. The first is that purchased by business units. It is often reasonable for the corporate modeler to work with distributions net of this reinsurance, at least initially—business units typically frown on corporate modelers criticizing their reinsurance decisions! The second is reinsurance purchased at the corporate level that covers a number of different business units. Catastrophe reinsurance is typically of this form. Optimizing the purchase of catastrophe reinsurance (and allocating its cost to business units) is a wonderful application of capital models, and corporate purchases should usually be modeled explicitly to facilitate it.

Counterparty credit risk should be handled by representing the probabilistic nature of ceded losses. What is supposed to be received as a reinsurance payout may or may not actually occur. Brehm et al. (2007) Section 6.1 is devoted to reinsurance receivables as a risk class.

Our InsCo example includes a 35 excess of 65 portfolio stop-loss contract with ceded losses of 35 occurring only in event 10. There is no counterparty risk. This example is discussed in Section 5.2.

2.4. Asset risk and Economic Scenario Generator

While the primary focus of the capital model is on how *insurance* liabilities impact the profitability and solvency of the firm, it may be desirable to represent contingencies on the asset side: first, to model assets as another source of volatility in the firm's financial position, and, second, to model how macroeconomic phenomena impact the insurance liabilities themselves. The key to modeling asset risk is a module that simulates interest rates and returns on various asset classes and possibly unemployment, inflation, and other socioeconomic indices. Such a module is generally referred to as an **Economic Scenario Generator**.

Brehm et al. (2007) Section 6.2 is devoted to investments. Daykin et al. (1993) Chapter 7 deals with inflation, and Chapter 8 is devoted to investments and asset/liability considerations such as the dynamics of reinvestment. Conning (2020) provides a comprehensive “basic guide to Economic Scenario Generators, with an emphasis on applications for the property/casualty insurance industry.” Also note that the American Academy of

Actuaries and the Society of Actuaries (SOA) have joined resources to manage the Economic Scenario Generators used in regulatory reserve and capital calculations (Society of Actuaries 2024). We are not using an Economic Scenario Generator in the InsCo example.

2.5. Capital structure

Capital structure, in corporate finance, refers to the division of liabilities corresponding to the sources of funds used to finance the firm. The basic division consists of shareholders' equity, (possibly) preferred stock, various types of debt, and possibly other liabilities. Reinsurance can also be considered part of the capital structure. See Mildenhall and Major (2022) for more on this perspective.

An important aspect of capital structure is its hierarchical nature insofar as there exists a pecking order in which the claimants providing these funds must be repaid in the course of liquidation. Insureds, providing premiums, come first. Common stockholders come last.

For an insurance entity, capital refers to the excess of assets over policyholder liabilities, whereas equity refers to the excess of assets over all other liabilities. Thus, debt can count as capital but not equity. For example, InsCo's assets consist of cash in the amount $a = 100$. The corresponding liabilities—the capital structure—consist of unearned premiums, a bond that InsCo had issued in the past and upon which it must pay a 3% annual coupon, and shareholder equity. See Table 2.2.

In the analysis presented in this monograph, it is usually sufficient to distinguish premiums P from capital Q (but adhere to Equation 1.1). However, it may be desirable to distinguish components of Q in reporting the financial fortunes of the firm. The hierarchy of claimants induces thresholds or benchmarks to distinguish degrees of financial distress or insolvency. This is elaborated upon in Chapter 3.

2.6. Accounting

Accounting is the language of business that defines the measurement of success or failure. It has its own vocabulary and grammar. You must have someone well versed in insurance accounting on the model development team. Whatever underlying representation for business cash flows the model uses must be translated into one or more standard accounting frameworks. This is material covered in the Casualty Actuarial Society's Exam 6. A good place to start is the detailed reading by Odomirok et al. (2020) or the earlier short introduction by Blanchard (2008).

Accounting is important for several reasons.

First, model output must be readily integrated with the firm's business plan. The plan is the key document orienting the goals of the company. The capital model is in-

tended to express and quantify the possible future variations from the plan that may emerge. In order to be useful and readily communicated to a variety of business actors in the firm, the model output needs to be expressed in the same terms as, and in a comparable format to, the plan. Companies often use their own management accounting conventions, usually slight variations on their reporting or regulatory standards. Therefore, the outputs must hew to the same accounting standards as the plan.

Second, accounting has real-world consequences. Insurers are subject to various types of valuation accounting standards, including:

1. Statutory or regulatory standards, such as US NAIC, EU Solvency II, APRA.
2. Financial reporting standards, such as GAAP and IFRS.
3. Rating agency standards, such as Standard and Poor's and AM Best's.

Failure to maintain an adequate financial position under any one of these standards could prove ruinous to the firm. This is elaborated on in Mildenhall and Major (2022) Section 8.3.

Emma et al. (2000) Chapter 7 opines that models should simultaneously represent at least cash (economic), statutory, reporting (GAAP, IFRS17), and tax accounting because “this is the only way to reflect the details of the interrelationships among constraints.” The model must model cash flows, which form the basis for the other accrual accounting models. In addition to representing external reporting standards, it is advisable to support your entity's management accounting, which is usually a slight variant of its reporting standard.

Figure 2.1 illustrates the relationship between three accounting views. For example, reserves are undiscounted under US GAAP and the market view of loss reserves may differ from management's.

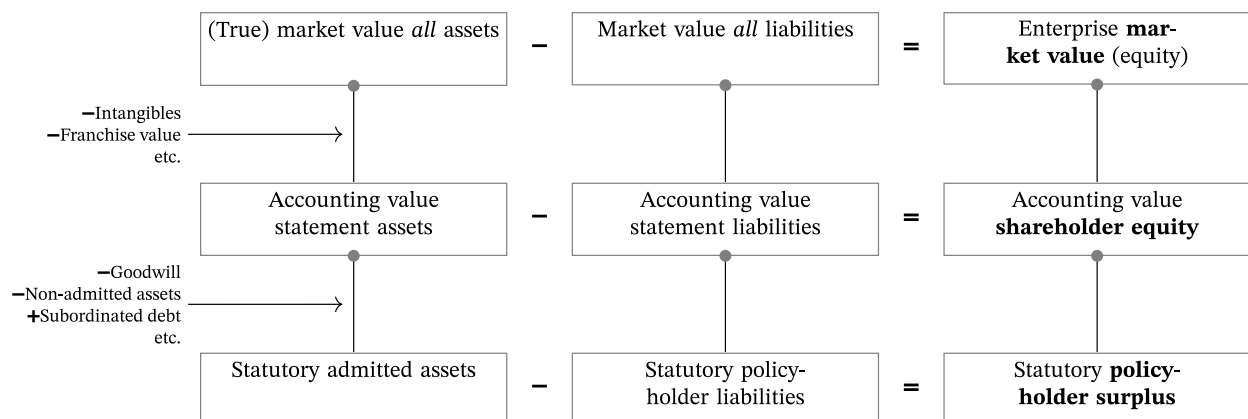


Figure 2.1. Capital, surplus, and equity under market value, GAAP, and statutory accounting views. From Mildenhall and Major (2022), used by permission.

Accounting produces income and balance sheet statements. The InsCo beginning period $t = 0$ balance sheet is shown in Table 2.2. Recall that the total plan premium is 52.2, received at the beginning of the policy period; expenses are excluded from this model and are assumed to be zero.

Table 2.2. InsCo's starting balance sheet.

Assets		Liabilities	
Cash	100.0	Unearned premium reserve	52.2
		Debt (bond)	20.0
		Shareholder equity	27.8

2.7. Pro formas

Pro forma financial statements are hypothetical financial statements about future states of the business. Whatever internal representation the model uses for outcomes, they must be translated into standard financial reports. We focus here on InsCo's balance sheet and income statements. Others are possible, including:

- Cash flow
- Sources and uses of funds
- Change in equity
- Comprehensive income
- Various statutory blanks, NAIC annual statement

For our one-period InsCo, the ending period shareholder value, if positive, is released to investors at time $t = 1$. For event 4 in Table 2.1, the income statement and balance sheet prior to paying losses and dividends are shown in Table 2.3 and Table 2.4. The balance sheet after payments is obviously 0 = 0 because InsCo is a one-period entity.

The Accounting module would produce 10 such pairs of financial statements.

Table 2.3. InsCo income statement, event 4.

Income		Expense	
Premiums	52.2	Loss & LAE	40.0
		Bond coupon	0.6
		Shareholder dividends	11.6

Table 2.4. InsCo ending balance sheet, event 4 prior to paying losses.

Assets		Liabilities	
Cash	100.0	Unearned premium reserve	0.0
		Loss reserves	40.0
		Debt (bond)	20.0
		Accrued interest	0.6
		Shareholder equity	39.4

Statutory reporting, done to satisfy regulatory authorities, is a deep and complex subject. Actuaries operating in the US should be familiar with the NAIC procedures manuals that can be found in Appendix V of Koca et al. (2023).

3. Capital adequacy and sources of risk

Too much of everything is just enough.

—Grateful Dead, “I Need a Miracle”

3.1. Capital adequacy

What quantity of assets is needed to back the liabilities the firm is taking on? The answer depends on who you ask. Insureds want more protection, meaning lower probability of insolvency and therefore higher assets. Regulators, representing the public, are aligned with the insureds. Investors want to minimize their own risk, so they want to invest less capital (for a given expected profit). Rating agencies are concerned with the counterparty (bond) creditworthiness and claims-paying ability of the firm, so they want to see an adequate level of lower-priority (equity, reinsurance) capital. The firm’s own management must adjudicate these competing demands in their own minds: on the one hand, they don’t want the firm to go under, but on the other hand, they need to deliver an acceptable rate of return to investors. They will have their own view of the ideal level of assets.

Even 100% collateralization of liabilities only theoretically guarantees 100% probability of solvency. In practice, insolvency can be caused by nonmodeled or inadequately modeled operational risks, such as fraud, rogue actors within the organization, management missteps, or bad-faith claim judgments in excess of policy limits. Capital adequacy is all about pegging an amount of assets deemed to be safe enough.

In practice it works like this:

1. Local regulatory authorities prescribe a minimum amount of *assets* and set reserving and valuation standards, which together imply a minimum *capital* requirement.
2. Rating agencies specify amounts that correspond to different rating levels (effectively, grades) given to the firm. Their models are generally based on statutory valuations but with adjustments, for example, for discount in loss reserves. The implied capital levels are usually higher than the regulatory minimums.
3. Management decides a target rating level (or in the EU, a ratio of capital to Solvency II required capital), balancing revenue and profit needs with what they believe their customers will tolerate. For example, personal lines markets are relatively insensitive to ratings, whereas large commercial markets require more highly rated insurers.

4. Management—with the help of the capital model—will look at the probabilities of adverse events causing the end-of-period assets to breach the above thresholds. They will then likely consider changing reinsurance, raising capital, or reducing exposure to keep these probabilities in control. In a startup situation, they might require an extra capital buffer, but a going concern rarely raises capital except for specific growth opportunities.

While insurance regulations and laws in the US are created and enforced by the individual states, the National Association of Insurance Commissioners (NAIC) develops model laws, regulations, and standards that most states follow. The NAIC created and maintains the Risk-Based Capital (RBC) framework, which is the primary method regulators use to assess capital adequacy. It specifies minimal capital requirements lest an insurer be put under watch or, worse, taken over by the regulators. NAIC (2024) provides background on RBC and links to related content.

Insurers in the European Union are bound by the Solvency II Directive (EIOPA 2009). This specifies a Solvency Capital Requirement to be calculated either by a standard model or by a firm's internal model. See IAA Risk Margin Working Group (2009), Floreani (2011), and Meyers (2019).

Rating agencies include S&P Global Ratings, AM Best, Moody's, and Fitch. They are primarily concerned with credit risk, i.e., assigning bonds a letter grade, but they also concern themselves with insurer claims paying ability, so there are two distinct types of rating. Rating agencies have their own models, often factor-based models similar to NAIC's RBC. Karakuyu et al. (2023) describe "S&P Global Ratings' methodology and assumptions for analyzing the risk-based capital (RBC) adequacy of insurers and reinsurers." Jakobsen et al. (2020) explain AM Best's Universal BCAR (Best's Capital Adequacy Ratio) model. Chan et al. (2024) "specifies Fitch Ratings' criteria for assigning new and monitoring existing international scale insurer Financial Strength (IFS) ratings." Readers should consult the agencies' websites for the most recent models, as these frameworks are periodically revised.

Larger insurance firms, especially publicly traded firms, tend to be organized into multiple legal entities, groups, and subsidiaries. Szkoda et al. (1995) state: "To properly analyze the financial condition of a company with subsidiaries, each subsidiary should be analyzed separately." Also, "understand debt obligations of company and parent company, look through on debt structure to determine if subsidiaries, etc., have sufficient cash flows to meet parent's obligations." Emma et al. (2000) say that either a consolidated model with internal cash flows between entities or entirely separate models (feeding each other) can be developed. Bear in mind that in stressed situations capital can become trapped.

Brehm et al. (2007) Section 2.3 compares regulatory and rating agency capital adequacy models. Eling and Holzmüller (2008) provide an overview and comparison of RBC requirements from the US, the EU, Switzerland, and New Zealand, although it is somewhat dated.

The firm's management will likely attend to a hierarchy of event possibilities (not necessarily distinct or in order):

- Missed earnings
- Spreads widen on bonds; stock price goes down
- Market value less than book value
- Bond ratings outlook change (watch)
- Bond ratings downgrade
- Claims paying ratings downgrade
- Being placed on regulatory watch
- Regulatory supervision, conservatorship or liquidation
- Loss of significant shareholder equity
- Loss of all shareholder equity
- Failure to pay bond interest
- Failure to fully redeem a bond
- Bankruptcy and reorganization

All these are highly correlated with the size of loss experienced in an accounting period, and all but missed earnings and market effects can be ameliorated by holding more capital.

Consider the InsCo example described in Table 3.1. If losses exceed 52.2 (premiums), then InsCo's investors will not receive any gain on their investment. This event has a probability of 30%. If losses are 80 or more ($52.2 + 27.8$ capital), then shareholders will be wiped out (lose all their investment); the probability is 10%. Sixty cents more and bondholders will not receive their promised coupons, and the company defaults and is put to the debt holders. If losses reach 100 (also a 10% chance), then bondholders will be

Table 3.1. InsCo critical loss levels.

Event	Loss Level	Probability
Zero profit	52.2	30%
Shareholders wiped out	80.0	10%
Failure to pay coupons	80.6	10%
Failure to redeem bonds	100.0	10%
99% average loss (TVaR)	100.0	10%

wiped out as well, suffering a 100% loss-given-default. Regulatory intervention thresholds might be defined by intermediate levels of loss and corresponding probabilities. Any of these might be used as VaR levels to define required assets. This is elaborated on in the next section.

3.2. Value at Risk and Tail Value at Risk

Various different risk measures can be used to calculate required capital. For a thorough discussion, see Section 3.6.3 of Mildenhall and Major (2022). Typically, the goal is to avoid a particular event with specified probability. An exceedance probability measure tells you the probability a firm becomes insolvent at a particular future time given an initial capital level. Inverting that relation determines the capital needed to attain the desired probability of safety. Thus there is a correspondence between risk measures and adequacy measures (Cherny and Madan 2009). Exceedance probability measures are common and lead to one or more **Value at Risk (VaR)** criteria: given a threshold probability, e.g., 99.5%, what size of loss will not be exceeded at that probability level?

VaR has a bit of a problem, however. It is not *subadditive*, which means there are situations where the VaR of the sum of two risks is greater than the sum of the two VaRs. Subadditivity is a good property for a risk measure because it means you manage risk from the bottom up: by managing the risk of each unit to be below its own threshold, you have managed the total risk to be below the sum of the thresholds. There are obvious implications for risk and capacity management. Theoretically, because VaR is not subadditive, it is not a good choice to manage risk at the unit level on up. However, this failure is more an academic quibble than a practical shortcoming. In many real-world situations VaR is subadditive, and it has the added benefits of being easy to understand, estimate, and back-test, and of always existing. In practice VaR is the risk measure of choice for insurance regulators around the world—with the exception of the Swiss.

Academics, the Swiss, and bankers in Basel 3 (also in Switzerland) posit a better alternative is **Tail Value at Risk (TVaR)**. TVaR asks, for a given probability level, what is the average loss in the worst events representing that accumulated probability? Defined in this way, TVaR is subadditive, a plus. On the downside, TVaR does not exist for thick-tailed distributions with no mean. It is also much harder to elicit and back-test than VaR. See Mildenhall and Major (2022) for an elaboration on this.

Another issue with TVaR is its precise definition. Historically, TVaR was called Expected Shortfall (ES) and Conditional Value at Risk (CVaR) by some authors. In addition, there are two subtly different measures called Tail Conditional Expectation (TCE) and Worst Conditional Expectation (WCE). In many cases these are all the same, but in many other realistic (discrete) cases they are not. See Chapter 4 of Mildenhall and Major (2022) for a thorough analysis of VaR and TVaR.

In many older texts, TVaR is defined as the average of losses above a specified VaR loss value. This coincides with the TCE definition for continuous random variables but not for discrete or mixed variables, and, in those cases, the resulting measure is not subadditive! More recent literature, led by McNeil et al. (2005), uses the definition we give. Formally,

$$\text{TVaR}_p(X) := \frac{1}{1-p} \int_p^1 \text{VaR}_s(X) ds = \frac{1}{1-p} \int_p^1 F^-(s) ds,$$

where $F^-(p)$ is the least inverse of the cumulative distribution function. In particular $\text{TVaR}_0(X) = E[X]$ and $\text{TVaR}_1(X)$ is defined to be $\sup(X)$, the largest possible outcome.

Here are two simple examples to help demystify subadditivity.

1. **Discrete example:** Consider a probability space with three outcomes, with probability 0.9, 0.05, and 0.05. Unit 1 has outcomes 0, 0, and 10, and unit 2 has 0, 10, and 0. The total outcomes are 0, 10, and 10. The $p = 0.9$ VaRs are 0 for each unit, but it is 10 in total. The corresponding TVaRs are 5 for each unit and 10 in total.
2. **Continuous example:** It is a curious fact that the sum of two standard exponential distributions fails to be subadditive at $p = 0.7$ (and neighboring p), so the failure can also apply for continuous variables. The unit VaRs are $-\log(0.3) = 1.204$ and the sum (a gamma with shape 2) has VaR equal to 2.439, which is greater than $2 \times 1.204 = 2.408$.

Mirroring the broader VaR-TVaR debate, your authors split in favor of VaR and TVaR. Your capital model must, however, employ *some* risk measure criteria to determine the amount of assets necessary to protect the firm against adverse loss experience. It can, of course, also cheat and select a (T)VaR threshold equating to actual capital held.

Our InsCo example uses assets $a = 100$, which can be characterized in various ways. One way is to say there is a $\text{TVaR}_{0.99}$ standard for the amount of assets needed. This is input “C” in Section 1.5. The other way says it is VaR_1 . Both result in $a = 100$, the maximum loss among the events, which is the asset value assumed previously. Notice two points. First, in practice the order of operations flows from accountants determining the amount of capital to modelers describing its adequacy (99% TVaR). Second, that $a = 100$ corresponds to full collateralization, and so it means there is no possibility of default. Default is not driving any results in this example! Nor should it in the preponderance of realistic models for well-rated companies. We might have chosen any VaR or TVaR level beyond 90% and obtained the same arithmetical results. We use TVaR to illustrate the algebra involved with a nontrivial asset risk measure.

In addition to specifying the model capital amount, it would be helpful to have some analytical capabilities to explain what drives the results. This is discussed in the next section.

3.3. Risk analysis

A main output of the Capital Adequacy module is an assessment of how often there is sufficient capital to avoid various undesirable outcomes. Natural follow-up questions revolve around explaining those results. Having drill-down and diagnostic capabilities is helpful, especially for complex models. These may include:

1. Statistics of selected variables
2. Identification of extreme events
3. Explaining what variables drive outputs
4. Evaluation of decision rules (e.g., reinsurance)
5. Risk/reward analysis
6. A battery of what-if questions

A systematic battery of what-if questions—What if trend is 3% higher than plan? What if hurricane frequency is actually 20% higher than the cat model assumptions?—can reveal where assumptions are particularly sensitive.

Even within the plan, there is an opportunity for sophisticated analysis. For example, machine learning could construct a decision tree to characterize simulation events with high losses.

At this point, there is a fork in the analytic road with two divergent paths that surprisingly rejoin to produce the same conclusion. The paths lead along a dynamic marginal impact analysis and a more static risk-adjusted probability analysis. Early work of Venter et al. (2006) saw the two paths, and more technical work of Delbaen (2000) actually knew the paths rejoined. Our work, Mildenhall and Major (2022), connected the dots more explicitly in an actuarial context.

The first form of risk analysis looks at the marginal impact on required assets of changes in the underlying portfolio. While unrealistic, a linear scaling of the unit loss distributions is often used, i.e., assets $a(X)$ become $a((1 + \epsilon)X)$. This is unrealistic because, e.g., doubling one's auto exposures homogeneously will not change X into $2X$. While it may approximately double the expected loss, the *shape* of the distribution will change: the standard deviation will less than double (Mildenhall 2017). Nonetheless, linear scaling is useful as a directional indicator. We will return to the issue of nonlinear scaling in Section 5.5.

Specifically, the **marginal approach** looks at the **gradient** (slope) of the risk measure: the change resulting from a small change in exposure. If $X = \sum_i X^i$ and $\phi(X)$ is a risk

measure, then define

$$\nabla^i \phi(X) = \frac{\partial}{\partial \epsilon} \phi(X + \epsilon X^i) = \lim_{\epsilon \downarrow 0} \frac{\phi(X + \epsilon X^i) - \phi(X)}{\epsilon}.$$

If ϕ is continuous and *homogeneous*, that is, $\phi(tX) = t\phi(X)$ for $t > 0$, then Euler's homogeneous function theorem tells us that

$$\sum_i \nabla^i \phi(X) = \phi(X),$$

i.e., the marginals consist of an allocation of the original measure (Mildenhall 2004).

In our example, the risk measure is $\phi(X) = \text{TVaR}_{0.99}(X)$. The gradient of this is

$$\nabla^i(\phi(X)) = \frac{1}{1-p} \int_p^1 \mathbb{E}[X^i \mid X = \text{VaR}_s(X)] ds,$$

often called the **coTVaR**. Those measures on the three units in our example refer again to the worst of the 10 losses, so the answers are 16, 20, and 64, respectively; see the last event in Table 2.1.

So far, the discussion refers to required *assets*, not *capital*. To calculate capital, we have to subtract premiums (Equation 1.1). This is detailed in Table 3.2. Ideally, *required* capital is calculated based on *required* premiums, but at this stage in our calculations we do not have those numbers.

Table 3.2. InsCo marginal analysis.

Item	Unit A	Unit B	Unit C	Portfolio
Required assets coTVaR, a	16.0	20.0	64.0	100.0
Expected loss, L	13.4	18.3	14.9	46.6
Plan premium	13.9	18.7	19.6	52.2
Required capital (plan)	2.1	1.3	44.4	47.8

Note that while TVaR submits to marginal analysis fairly well, VaR needs special handling. Say we were concerned with $\text{VaR}_{0.85} = 65$. This occurs solidly in event 9 (Table 2.1). A straightforward marginal analysis breaks this down to Table 3.3.

Table 3.3. InsCo marginal analysis of $\text{VaR}_{0.85}$.

Unit A	Unit B	Unit C	Portfolio
17.0	8.0	40.0	65.0

Yet, for both the lower event 8 and higher event 10, line B's contribution is higher. Imagine a more detailed simulation where each event is only worth 1/10,000 probability. Looking at just that specific event, defining the VaR is inviting a good deal of instability. Another run with different random numbers will likely produce materially different results.

There are two ways out of this problem. One is to estimate the conditional expected loss values described in Section 4.3. The other is to use a more robust replacement for VaR known as **Range Value at Risk (RVaR)** (Cont et al. 2010). This is the average loss occurring between two probability levels. For example, instead of using $\text{VaR}_{0.99}$, you might use $\text{RVaR}_{0.985,0.995}$.

A sign of problems of the marginal approach is the need to stipulate that epsilon decreases to zero. When there are ties in the outcome distribution, the risk measure is not actually differentiable! It has a cusp, and depending on whether you increase or decrease volume in a given line, you get different answers. The risk measure is analogous to the absolute value function at zero, where the left and right derivatives are different. (We will see this haunt us in Section 5.5.) To ensure the risk measure is differentiable, and a marginal approach is really valid, it is enough to require that all outcomes of the total distribution be distinct. In that case, for a small enough change in any unit, there is no reordering of the total outcomes and the risk measure is differentiable. We will implement that requirement in the next chapter.

The second form of risk analysis, again following Venter et al. (2006), looks for a set of risk-adjusted probabilities that reproduce the results of the risk measure. With these adjusted probabilities in hand, it is natural to use the risk-adjusted expected value of each unit as an allocation *of risk*. This method is the basis of the *Natural Allocation* described in Mildenhall and Major (2022) and picked up in Chapter 4 to allocate *premium*.

3.4. Algorithms

3.4.1. Algorithm to evaluate expected loss for discrete random variables

The algorithm in this subsection is very basic. We present it to establish an approach to working with discrete random variables that we generalize in subsequent chapters.

We present an algorithm to compute $E[X]$ in two ways, based on

$$E[X] = \int_0^{\infty} x dF(x) = \int_0^{\infty} S(x) dx. \quad (3.1)$$

Algorithm Input: X is a discrete random variable, taking finitely many values $X_j \geq 0$, and $p_j = P(X = x_j)$.

Follow these steps to evaluate Equation 3.1.

Algorithm Steps

- (1) **Pad** the input by adding a zero outcome $X = 0$ with probability 0.
- (2) **Group by** X_j and sum the corresponding p_j .
- (3) **Sort** events by outcome X_j into ascending order. Relabel events $X_0 < X_1 < \dots < X_m$ and probabilities p_0, \dots, p_m .
- (4) **Decumulate** probabilities to compute the survival function $S_j := S(X_j)$ using $S_0 = 1 - p_0$ and $S_j = S_{j-1} - p_j, j > 0$.
- (5) **Difference** the outcomes to compute $\Delta X_j = X_{j+1} - X_j, j = 0, \dots, m - 1$.
- (6) **Outcome-probability sum-product:**

$$E[X] = \int_0^\infty x dF(x) = \sum_{j=1}^m X_j p_j. \quad (3.2)$$

- (7) **Survival function sum-product:**

$$E[X] = \int_0^\infty S(x) dx = \sum_{j=0}^{m-1} S_j \Delta X_j. \quad (3.3)$$

Comments

- (a) Step (1) treats 0 as special because the second integral in Equation 3.1 starts at $X = 0$. Step (1) allows us to systematically deal with any discrete data. It adds a new outcome row only when the smallest observation is > 0 .
- (b) After Step (3), the X_j are distinct; they are in ascending order, $X_0 = 0$, and $p_j = P(X = X_j)$.
- (c) In Step (4), $S_m = P(X > X_m) = 0$ since X_m is the maximum value of X .
- (d) The forward difference ΔX computed in Step (5) replaces dx in various formulas. Since it is a forward difference, ΔX_m is undefined. It is also unneeded.
- (e) In Step (6), the sum starts at $j = 1$ because $X_0 = 0$. Notice that $P(X = X_j) = S_{j-1} - S_j$ is the negative backward difference of S .
- (f) Note the index shift between Equation 3.2 and Equation 3.3.
- (g) Both Equation 3.2 and Equation 3.3 are exact evaluations. The approximation occurs when the underlying distribution being modeled is replaced with the discrete sample given by X .

3.4.2. Algorithm to evaluate VaR for a discrete distribution

Algorithm Input: X is a discrete random variable, taking N equally likely values $X_j \geq 0$, $j = 0, \dots, N - 1$. Probability level p .

Follow these steps to determine $\text{VaR}_p(X)$.

Algorithm Steps

- (1) **Sort** outcomes into ascending order $X_0 < \dots < X_{N-1}$.
- (2) **Calculate** $n = \lfloor pN \rfloor$, the greatest integer less than or equal to pN .
- (3) **Return** $\text{VaR}_p(X) := X_n$ (since the indexing starts at 0).

These steps compute the least quantile at level p . A p quantile is any value x so that $P(X < x) \leq p \leq P(X \leq x)$ and VaR_p is the *smallest* such x . If you draw a graph of the distribution function and allow the vertical lines joining jumps to be part of the graph, then this algorithm is just inverting the distribution function by finding the leftmost x at which the graph reaches height p ; see Chapter 4 of Mildenhall and Major (2022).

3.4.3. Algorithm to evaluate TVaR for a discrete distribution

Algorithm Input: X is a discrete random variable, taking N equally likely values $X_j \geq 0$, $j = 0, \dots, N - 1$. Probability level p .

Follow these steps to determine $\text{TVaR}_p(X)$.

Algorithm Steps

- (1) **Sort** outcomes into ascending order $X_0 < \dots < X_{N-1}$.
- (2) **Find** n so that $n \leq pN < (n + 1)$.
- (3) **If** $n + 1 = N$, **then** $\text{TVaR}_p(X) := X_{N-1}$ is the largest observation; exit.
- (4) **Else** $n < N - 1$ and continue.
- (5) **Compute** $T_1 := X_{n+1} + \dots + X_{N-1}$.
- (6) **Compute** $T_2 := ((n + 1) - pN)x_n$.
- (7) **Compute** $\text{TVaR}_p(X) := (1 - p)^{-1}(T_1 + T_2)/N$.

These steps compute the average of the largest $N(1 - p)$ observations. Step (6) adds a pro rata portion of the $\lfloor N(1 - p) \rfloor$ largest observation when $N(1 - p)$ is not an integer.

4. Pricing and allocation

This chapter is about the basic theory of allocating premium and capital based on the Natural Allocation of Spectral Risk Measures (SRMs). Here, we present the logic in an informal fashion. Think of it as the CliffsNotes or Schaum's Outlines version. Topics include:

- The one-year InsCo model
- The relationship between assets, capital, and premium
- Capital risk measures versus pricing risk measures
- Portfolio pricing
- Layer pricing
- Event pricing
- Unit pricing
- The constant cost of capital assumption.

Allocating capital *per se* is addressed in Section 8.3.

The theory presented here revolves around a one-period model. In reality, insurance claims often evolve over time with multiple payments spanning more than one accounting period. The output of the Business Operations module will undoubtedly reflect this. In order to apply the theory, we must collapse the multiperiod view into a single-period equivalent. There are two ways to do this. One, we can represent paid claims in a unit and carry a one-year change in reserves as a separate but parallel unit. Two, we can operate entirely in ultimate claims, discounting them back to time zero. To do this, we lean on the principles espoused in Wüthrich et al. (2010):

- If interest rates are not being modeled (you do not have an Economic Scenario Generator), use the latest available risk-free zero-coupon bond yield curve to discount future simulated paid claims to the baseline time in question.
- If an Economic Scenario Generator is available to generate random interest rates, use the generated zero-coupon bond yield curve in each event to discount payments in the same event.

Using risk-free rates to discount losses is somewhat controversial. Fairley (1979) argued for a risk-free rate because “policyholders do not share in [the insurer’s] investment

risks....” Cummins (1990) argued for using the market return on the company’s actual portfolio because adding investment risk increases the probability of default for policyholders and therefore changes their expected loss recoveries, i.e., they *do* share in the investment risk. However, note that the insolvency priority order of payments puts policyholders first in line to be paid and shareholders last (i.e., first in line for losing money). Investment risk is not shared equally—if we want to follow Cummins’s line of reasoning, the correct discount rate for losses is nearer risk-free than it is the full investment yield. The risk-free rate is probably not a bad approximation.

An alternative theory is embedded in IFRS17 rules (Caramagno et al. 2021). There, a discount rate for future liability cash flows is the sum of a risk-free rate and a liquidity premium. Since insurance contracts cannot be traded continuously in a liquid market, investors must therefore demand a liquidity premium. Typically this is around 50 basis points. The IFRS17 rate proxies an entity independent market rate.

For more complex situations where cash flows depend more directly on financial outcomes (as is the case in much of life insurance), refer to Wüthrich et al. (2010)—but not the parts about risk loads; their approach is not the same as the SRM approach presented here!

4.1. Pricing the whole

Pricing applications of capital models typically proceed from the top down (Bühlmann 1985) and start by determining the premium needed for the whole portfolio to meet corporate return objectives. Total needed return depends on the amount of capital and the required return on that capital. Following the market structure laid out for InsCo (Figure 1.4), the amount is determined by regulators or rating agencies using a capital risk measure. The return is estimated using a separate analysis that will be used to calibrate the pricing risk measure.

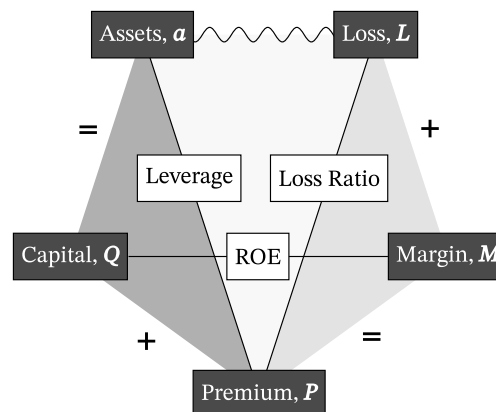


Figure 4.1. The insurance pentagon.

A stock insurer uses a peer analysis of valuation against relevant financial metrics to estimate what it believes its investors require as a return on equity to support the current or desired share price. This return is combined with the capital structure (known debt costs) to determine a weighted average cost of capital across debt and equity, producing an implied needed total income. Income is then adjusted for non-underwriting income and taxes to determine needed income from insurance operations. It is this amount that is handed to the capital modelers to be allocated to individual units (Section 4.2). The analysis below ignores non-underwriting income and taxes and proceeds directly from the implied weighted average cost of capital. Mutual companies substitute their own analysis, reflecting their founding charter and mission, and their lack of access to equity financing, to determine an analogous cost of capital.

Say the analysis determines that investors require a 15% return on capital for their investment in InsCo. The portfolio's premium, margin, capital, assets, and expected losses have a simple relationship with rate of return. The time value of money is ignored, i.e., the risk-free rate is zero. Here we analyze only the gross losses, leaving the analysis of reinsurance for later.

Define the expected loss $L = E[X \wedge a]$. L is slightly less than the promised loss $E[X]$ because of the possibility of default; in practice the adjustment is so small you can often ignore it. Premium can be decomposed into expected loss plus margin:

$$P = L + M. \quad (4.1)$$

Equation 1.1, the funding equation, says $a = P + Q$ giving the split of assets between policyholder premium and investor capital. Policyholders in aggregate have been promised X and investors the balance of funds $(a - X)^+$. If $X > a$ policyholders are paid a and investors nothing. No matter the loss outcome, all ending assets are paid in losses or returned to investors:

$$a \equiv X \wedge a + (a - X)^+.$$

Investors *expect* to receive

$$E[(a - X)^+] = a - L = (P + Q) - L = (L + M) + Q - L = M + Q,$$

and therefore their expected return on investment is $\iota = M/Q$.

In Section 10.5 of Mildenhall and Major (2022), we introduce the insurance pentagon, Figure 4.1, which displays key insurance pricing variables and the relationships between them. There are five monetary variables: premium, loss, margin, assets, and capital, and three ratios describing leverage, loss ratio, and return (ROE).

Table 4.1 displays a number of relationships between the eight key insurance variables that we use repeatedly: you should become familiar with them. The table uses Greek letters for the ratios and Roman for monetary amounts.

Table 4.1. The eight pentagon variables and relationships between them and with other insurance statistics.

Ref	Variable or Relationship	Interpretation
Monetary amounts		
1	L	Expected loss
2	P	Premium
3	M	Margin
4	a	Total assets
5	Q	Capital
Related amounts		
	$a - L$	The unfunded liability above expected loss, funded by margin and capital
Monetary identities		
Prem	$P = L + M$	Premium is expected loss plus margin
Fund	$a = P + Q$	Funding equation: premium and capital only source of $t = 0$ assets
Ratios		
6	$\lambda = L/P$	Loss ratio
7	$\gamma = P/a$	Premium to asset leverage (gamma = g for leveraGe)
8	$\iota = M/Q$	Expected return on capital (investor) or cost of capital (insured)
Related ratios		
	$\nu = 1/(1 + \iota)$	Risk discount factor, analog of $\nu = 1/(1 + i)$
	$\delta = \iota/(1 + \iota)$	Risk discount rate, analog of paying interest at $t = 0$
7a	$P/Q = \gamma/(1 - \gamma)$	Premium to capital leverage ratio, divide top/btm by a , used Fund
Ratio identities		
	$\delta = \nu$	Analog of $d = iv$ in theory of interest
Disc	$1 = \nu + \delta$	Analog to $1 = v + d$
Capital identities		
Q	$Q = \nu(a - L)$	Q is $t = 0$ price paid for asset with expected $t = 1$ return $a - L$
8a	$\iota = (P - L)/(a - P)$	Substitute Prem and Fund into (8)
	$\delta = (P - L)/(a - L)$	Substitute Prem and Fund into (8)
Premium identities		
P1	$P = \nu L + \delta a$	Premium = weighted average of expected loss and maximum loss, a
P2	$P = a - \nu(a - L)$	Premium = assets not funded by investor capital
P3	$P = L + \delta(a - L)$	Premium = expected loss plus share of unfunded liability
P3a	$P = L + \iota Q$	As P3 but expressed in terms of cost of capital

Premium (like discount) is received at $t = 0$, hence the relevance of δ . Income (like interest) is earned and paid at $t = 1$ and measured with ι . Discount is easier to work with as a measure of return because it is always between 0 and 1: $\delta = 1$ corresponds to an infinite return.

Pricing is the process of splitting the **unfunded liability** $a - L$ into premium margin and capital parts funded by the policyholder and investor. The risk discount factors tell you how to accomplish this.

The relationship between premium and loss can be expressed using the loss ratio or its reciprocal, the premium multiplier. The latter is sometimes used in catastrophe pricing, when loss ratios seem embarrassingly low. Likewise, leverage could be expressed as premium to capital rather than as premium to assets. The point is there is a measure of leverage, a measure of margin to volume, and a measure of return-cost to capital.

Given required expected return from a top-down approach analysis, the standard pentagon equations in Table 4.1 show that knowing ι , L , and a determines total premium. It can be expressed in multiple ways, the most helpful of which are

$$P = L + \delta(a - L) = L + \iota Q = \nu L + \delta a. \quad (4.2)$$

The first two expressions make it clear that premium equals expected loss plus cost of capital. The third will reappear again and again in this monograph.

Example: TVaR Capital Risk Measure. Before considering allocation, let's look at total pricing with a TVaR capital measure, $a = \text{TVaR}_p(X)$. The quantity $a - L = \text{TVaR}_p(X) - E[X]$ is usually referred to as $\text{XTVaR}_p(X)$, for excess TVaR above the mean. Total premium, from Equation 4.2 becomes

$$\begin{aligned} P &= L + \delta(a - L) \\ &= L + \delta \text{XTVaR}_p(X) \\ &= L + \iota Q. \end{aligned}$$

Example: InsCo. Recall from Table 3.2 that the portfolio plan premium is 52.2. When $\iota = 0.15$, we have $\nu = 0.86956$ and $\delta = 0.13043$, resulting in a required portfolio premium of $46.6 + 0.13043 \cdot (100 - 46.6) = 53.565$. The premium required to achieve a 15% return is 53.6, or 1.4 higher than plan. The required margin is 7.0, also 1.4 higher than the plan margin. This difference between plan and required profitability (negative here) is sometimes called **economic value added** or EVA. It represents the amount of profit over and above what is required to satisfy investors. In Section 5.5 we see if we can improve the situation. In the rest of this chapter, we address how to allocate the required portfolio premium and margin to the units.

The pentagon contains eight variables and five relationships (funding, premium, loss ratio, leverage, and return), leaving three free variables, i.e., most of the time knowing three of the eight variables is enough to determine the rest. This sometimes fails, e.g., knowing the three ratios does not determine volume.

You should become familiar with the underlying algebra of Table 4.1, not memorize it, but be able to figure it out quickly. Here are some examples.

- Using $\nu + \delta = 1$ and starting from $P = \nu L + \delta a$ gives $P = (1 - \delta)L + \delta a = L + \delta(a - L)$. Since we know $P = L + M$, this implies $M = \delta(a - L)$.
- Alternatively, $P = \nu L + (1 - \nu)a = a - \nu(a - L)$. Now, since we know $a = P + Q$ so $P = a - Q$, this implies $Q = \nu(a - L)$.
- Hence return $\iota = M/Q = \delta/\nu$.

Make sure you are happy with P3 and P3a. These are used repeatedly below.

4.2. Allocation to units: Overview

Introducing the Natural Allocation. The Natural Allocation (NA) is a way to allocate total premium or margin down to individual units. The method is natural because it relies on so few assumptions, is easy to compute, and gives a reasonable range of interpretable answers. It also agrees with the marginal and risk-adjusted probability methods when they are well defined—a huge conceptual advantage. The NA works by taking the problem of pricing the whole risk and dividing it into a series of simpler problems involving pricing individual layers of assets and then solving those simpler problems. Layers are easier to price because they are binary, involving no loss or a full loss, and are completely characterized by their probability of a loss. The NA is described in the next six sections.

- Section 4.3 describes four adjustments that must be made to raw simulated output before starting. These are technical, computational adjustments.
- Section 4.4 describes how the method works in total. It concludes by showing that the NA depends on the selection of a distortion function to determine the return to each layer of assets.
- Section 4.5 explains how to move from a total price for each layer to a price by layer by unit. Summing these allocations up across layers gives the unit NA.
- To this point, we have not introduced any distortion functions, so Section 4.6 introduces five families. These essentially span the full range of risk appetites. We compute the NA for the Wang distortion, which corresponds to a middle-of-the-road risk appetite.
- Section 4.7 describes how the NA works for the constant cost of capital distortion (CCoC), which assigns the same return to each layer. The CCoC corresponds to a very

extreme tail-risk averse risk appetite. Unfortunately, while the CCoC is an unrealistic assumption, it is the basis for what we call the industry standard approach. It is the root cause of many problems encountered in practical allocation work. We explain why and how it came to be accepted.

- Finally, Section 4.8 tries to head off objections to the NA caused by interpreting scenarios defined by outcome loss amount as events. If this idea has not occurred to you before already, we recommend skipping it! However, someone invariably brings it up, and you need to be able to explain why it is not a problem, so you do need to read it eventually.

4.3. Four technical adjustments

In order to make the algebra work out in subsequent sections, we need to adjust our events so that four simplifying conditions hold. These alter the form of the simulation data, but not the underlying meaning. We adjust our simulated loss outcomes so that

1. Loss outcomes are sorted by total loss, from lowest to highest.
2. The first loss is zero.
3. Losses are capped at the available assets.
4. Portfolio losses are unique by event.

Zero Loss. If, in our table of loss probabilities, the first, lowest, portfolio loss X_1 is greater than zero (and this is true in Table 2.1), then it is computationally simpler to introduce another row with $X_0 = 0$ and $p_0 = 0$. This does not change the results because $p_0 = 0$. If $X_1 = 0$ already, then no such adjustment is necessary. However, in that case, we do renumber so that the first event is called $k = 0$. The reason for this will become apparent in the next section; layer size and exceedance probability must be properly accounted for.

Capped Losses. When there is zero probability of insolvency, allocating the portfolio's expected loss $E[X]$ to unit i is easy—it's just the unit's expected loss $E[X^i]$. When assets are lower and there is a nonzero probability of default $X > a$, we need to consider what actually happens to claim payouts. The most common rule, and the one we assume, is **equal priority**. Payments to unit i are given a common haircut:

$$X^i(a) := X^i \frac{X \wedge a}{X} = \begin{cases} X^i & X \leq a \\ X^i \frac{a}{X} & X > a. \end{cases}$$

This expresses payments to unit i as a constant pro rata (haircut) proportion $(X \wedge a) / X$ of the contractually promised payment X^i . The factor is the same for all units, hence *equal* priority. We assume the appropriate substitutions have been made in our loss table:

$$\begin{aligned} X^i &\leftarrow X^i(a), \\ X &\leftarrow X \wedge a. \end{aligned}$$

It is important to realize that once this adjustment has been made we can **ignore default**. Default has not disappeared; by replacing promised loss payments X^i with actual $X^i(a)$, we have “baked in” default. Going forward we can therefore ignore default as a separate phenomenon.

Unique Losses. Consider a particular outcome $X = X_k$. If there is only one event with $X = X_k$, then there is only one set of unit losses $\{X_k^1, \dots, X_k^m\}$ whose sum is X_k . If there are multiple events with $X = X_k$ but the same set of $\{X_k^i\}$, then we can, with no loss of information, collapse them all (summing the probabilities) and consider the result to be one event.

However, if there are two or more distinctly different events resulting in $X = X_k$, then the above logic cannot be applied. This is true in Table 2.1. In this case marginal calculations do not make sense, and when, in Section 4.4, we go to distort probabilities, there will be an ambiguity in the ordering of the events that will be material to the calculations. We need to resolve that ambiguity, and we can do it by using the *conditional expected loss* of the units. Define

$$\kappa^i(x) := E[X^i \mid X = x].$$

Then collapse the events sharing $X = X_k$, replacing the particular X^i values with their conditional expectation $\kappa^i(X_k)$ and summing the probabilities into the newly defined p_k . We appeal to a symmetry argument: within the set of events resulting in $X = X_k$, beyond their relative probabilities, there is no reason to prioritize one over another. Henceforth, we will assume the substitutions

$$X_k^i \leftarrow \kappa^i(X_k) \tag{4.3}$$

have been made in our table of outcomes. Notice that Equation 4.3 always applies when losses exceed assets. The substitution Equation 4.3 relies on equal priority in the sense that the recovery to the collapsed event equals the conditional expectation of the recov-

eries to its components:

$$\begin{aligned} \mathbb{E} \left[X^i \frac{X \wedge a}{X} \mid X = x \right] &= \mathbb{E} [X^i \mid X = x] \frac{X \wedge a}{X} \\ &= \kappa^i(X) \frac{X \wedge a}{X}. \end{aligned}$$

Thus, equal priority is essential to make the approach work. Luckily, equal priority is universal in insurance regulation, even though other approaches are possible in theory (Mildenhall and Major (2022), Section 12.3).

In our example, notice in Table 2.1 that the events 4, 5, 6, and 7 all have $X = 40$. Taking the p -weighted averages of each unit's losses across those events, collapsing the events down to one event, renumbering, and inserting the zero row, we get Table 4.2. The table has also added a row 0 with 0 loss and probability. There is no need for a default adjustment because $a = 100$, sufficient to pay the maximum loss.

Table 4.2. InsCo loss probability distribution by unit with duplicate X averaged out and zero row inserted. Index now called k rather than j .

k	p	Unit A X_k^1	Unit B X_k^2	Unit C X_k^3	Portfolio X_k
0	0	0	0	0	0
1	0.1	15	7	0	22
2	0.1	15	13	0	28
3	0.1	5	20	11	36
4	0.4	10	24	6	40
5	0.1	26	19	10	55
6	0.1	17	8	40	65
7	0.1	16	20	64	100
$\mathbb{E}[X]$		13.4	18.3	14.9	46.6
Plan premium		13.9	18.7	19.6	52.2

4.4. Pricing layers in total

Layer pricing assumes the required return associated with each individual dollar of assets is given by a function of its risk of being consumed for claim payments.

The NA is based on the principle of *layer pricing*, introduced formally by Wang (1996) and articulated by Venter (1991), Bodoff (2007), and others.

Suppose the total assets a are funded by premium and the sale of the residual value of separate **layers** or **tranches**. These layers are analogous to collateralized aggregate reinsurance, as we shall now explain. Reinsurance in our one-period model is a transaction where ceded premium is paid to the reinsurer at $t = 0$ and ceded losses are recovered from the reinsurer at $t = 1$. Let's turn that around to make the role of reinsurance as

capital more transparent. InsCo (or its regulator) demands *collateralized* reinsurance. Consider a layer of width ΔX_k . The layer has its own funding equation and needs to be supported by assets of ΔX_k in total. At $t = 0$ the counterparty remits $\nu(\Delta X_k - L_k)$ and the insurer $L_k + \delta(\Delta X_k - L_k)$ into the collateral account, where L_k is expected loss in the layer. Since $\nu + \delta = 1$, the layer is fully collateralized. At $t = 1$ the layer loss becomes known and is paid in full from collateral. Any amounts remaining are remitted back to the counterparty. This is reinsurance, but with the transactions reversed to reveal the capital financing function.

To be concrete, say we understand the distribution of portfolio losses X through a simulation. Say the annual aggregate results

$$X \in \{X_0 < X_1 < \dots < X_n = a\}$$

and each outcome has its probability $\{p_0, p_1, \dots, p_n\}$.

We tranche total assets a into a stack of fully collateralized aggregate reinsurance layers. Each layer is a half-open interval, $k = 0, \dots, n-1$, covering losses $X \in (X_k, X_{k+1}]$, i.e., for a loss of X it pays

$$\begin{cases} 0 & X \leq X_k \\ X - X_k & X_k < X \leq X_{k+1} \\ \Delta X_k & X_{k+1} < X. \end{cases}$$

However, there are no partial payments because X_k are the only possible outcomes, so the cover reduces to one paying nothing when $X \leq X_k$ and ΔX_k as soon as $X > X_k$.

For our InsCo numerical example, $X \in \{0, 22, 28, 36, 40, 55, 65, 100\}$ and $n = 7$. So, we have seven layers $(X_k, X_{k+1}]$ that we imagine are funded separately, each by a combination of investor capital (security market purchase price of the returned collateral) and insured's premiums. The probability that losses exactly equal X_k is p_k . The probability that layer k is used for loss payments (rather than handed back to investors) is the *exceedance probability* or *survival function* (hence notation S):

$$S_k = S(X_k) := \mathbf{P}\{X > X_k\} = \sum_{j=k+1}^n p_j.$$

The numbers for our example are laid out in Table 4.3.

How is layer k to be funded? Analogous to our portfolio funding Equation 1.1, we have the **layer funding equation**:

$$\Delta X_k = P_k + Q_k. \quad (4.4)$$

Table 4.3. InsCo layer loss details and survival probabilities.

k	p	Unit A X_k^1	Unit B X_k^2	Unit C X_k^3	Portfolio X_k	Layer Size ΔX_k	Exceedance Pr S_k
0	0	0	0	0	0	22	1
1	0.1	15	7	0	22	6	0.9
2	0.1	15	13	0	28	8	0.8
3	0.1	5	20	11	36	4	0.7
4	0.4	10	24	6	40	15	0.3
5	0.1	26	19	10	55	10	0.2
6	0.1	17	8	40	65	35	0.1
7	0.1	16	20	64	100	0	0

While the portfolio funding equation is self-evident, is it necessary to constrain a layer's capital to $Q_k = \Delta X_k - P_k$, or can it be something else? It is easy to see it cannot *provided* $P = \sum_k P_k$. Let's assume that (we'll justify it later) and see why. First, note that $Q_k \geq \Delta X_k - P_k$ for each k , otherwise there could be a shortfall in loss payments on the layer. This is just the statement that each layer must be fully collateralized. Now, if the inequality is strict for any k , we get $Q = \sum_k Q_k > \sum_k \Delta X_k - P_k = a - P = Q$, which is a contradiction. Hence $Q_k = \Delta X_k - P_k$.

We still need to justify $P = \sum_k P_k$. You might think that the aggregate premium could be less than the sum of the premiums by layer because of potential diversification opportunities between the layers. However, in a sense there are no diversification opportunities between the layers because they are **comonotonic**: they are all increasing functions of the total loss X . Some, but not all, pricing functionals are **comonotonic additive**, meaning that the price of the sum of comonotonic risks equals the sum of the prices. To proceed, we need to use a comonotonic additive pricing functional. This is not a problem, as the functionals we are proposing to use are comonotonic additive. Thus, we may assume the layer funding Equation 4.4 holds.

We have reduced the pricing problem to pricing a fully collateralized layer. Here again we can apply our general rules: if we know the expected loss and *return by layer*, we know the premium. S_k is the probability that layer k is needed to pay claims and so $L_k = S_k \Delta X_k$ is the expected value of layer k 's contribution to the total portfolio expected claim payments L . For the return, we need to make one last assumption: S_k is the only thing that investors deem relevant to pricing layer k . In layer k , the investors get ΔX_k back with probability $1 - S_k$ and nothing with probability S_k . Layer k 's payoff as a fraction of its limit ΔX_k is a Bernoulli random variable with one parameter: $1 - S_k$. We assume that its probability distribution, determined by S_k , is all that matters. This restriction is known as law invariance.

A risk measure $\rho(X)$ is **law invariant** if and only if it is a function of the distribution F of X . That is, if X and Y have the same distribution function, then $\rho(X) = \rho(Y)$. Law invariant risk measures do not depend on the cause of loss or the particulars of the event generating the loss. VaR, TVaR, and standard deviation are examples of law invariant risk measures. A measure like $\rho(X) = E[X] + \lambda \text{COV}(X, Y)$, which depends on a state variable Y , is not.

Law invariance may seem a strange assumption for a pricing risk measure, especially in the light of theories like the capital asset pricing model where systematic or nondiversifiable risk is considered (these are not law invariant). The rationale is that insurance risk is diversifiable and has no analog of market risk. The motivation for law invariance in the *capital* risk measure stems from a regulator's desire for an objective default probability measure. An entity's risk of insolvency depends only on the distribution of its future change in surplus—the cause of loss is irrelevant. Law invariance enables risk to be estimated statistically, from an empirical sample or a model fit to X . Law invariance is sometimes called **objectivity** for these reasons. It allows regulators to calibrate their models using only losses. While the justification for law invariance in the *pricing* measure is admittedly weaker, we find it eminently useful.

Suppose we know (or, being actuaries, assume) a function $g(s)$ that maps attachment probability S_k to rate on line (premium to limit ratio) allowing us to write

$$P_k = g(S_k) \Delta X_k.$$

The function $g(s)$ is known as a **distortion function**.

Before moving on, here is a summary of the key assumptions that we have made and how they are used:

1. Equal priority in default—collapse equal outcomes to one.
2. Comonotonic additive pricing risk measure—funding equation by layer.
3. Law invariant pricing risk measure—price layer using a distortion function.
4. Existence of a rate on line distortion function g for pricing Bernoulli layers.

Distortion functions must have certain properties if the layer prices are to be rational. In particular, $g(0) = 0$, $g(1) = 1$, and g must be increasing and concave (decreasing slope). Violations of these lead to irrational pricing behavior, as defined and explained in Mildenhall and Major (2022), Chapter 10.

The logic for pricing layer k is exactly the same as what was used to derive Equation 4.2 and gives

$$P_k = g(S) \Delta X_k = \nu_k S_k \Delta X_k + \delta_k \Delta X_k, \quad (4.5)$$

where the return ι_k and risk discount factors ν_k and δ_k are *functions of S* (hence the subscripts). So, if we knew the required return ι_k associated with layer k , we could determine its premium and capital. Summing across all layers we would then know both total premium and total capital, and they would sum to total assets a . Computing the total premium

$$P = \sum_j g(S_j) \Delta X_j$$

in this way is known as a **Spectral Risk Measure (SRM)**. SRMs are *coherent* risk measures in the sense of Artzner et al. (1999). They are also *law invariant* (of course) and *comonotonic additive*, and in fact are the only risk measures that have all these properties.

Let's look more closely at Equation 4.5. From our setup, the largest possible loss and $\sum \Delta X_k = a$. Using that, recall the two ways of computing the mean (integration by parts):

$$E[X] = \sum_{k=0}^n X_k p_k = \sum_{k=0}^{n-1} \Delta X_k S_k.$$

The right-hand term is a sum over expected losses in each layer k . This formula is illustrated in Figure 4.2 by the blue area and explained in the context of the Lee diagram in Chapter 3 of Mildenhall and Major (2022). Now, if each layer of capital costs $g(S_j)$, we can use integration by parts to get

$$\begin{aligned} P &= \sum_{k=0}^{n-1} g(S_k) \Delta X_k \\ &= \sum_{k=0}^{n-1} g(S(X_k)) (X_{k+1} - X_k) \\ &= \sum_{k=0}^n (g(S_{k-1}) - g(S_k)) X_k \\ &= \sum_{k=0}^n q_k X_k, \end{aligned} \tag{4.6}$$

where $g(S_{-1}) := 1$. Refer to Figure 4.2. The size of each vertical step is ΔX_k , and the left-to-right distance from each blue (respectively, orange) vertical step to the $p = 1$ axis is S_k (respectively, $g(S_k)$). The area under the blue line is the expected loss. The area under the orange line is the premium, corresponding to either expression in Equation 4.6. The difference is the margin, shown in orange shading in the second panel. The complement of premium is capital, above the orange line, shaded in light blue. This view of the relationship between capital, margin, and loss is dramatically different from the conventional view in the third panel.

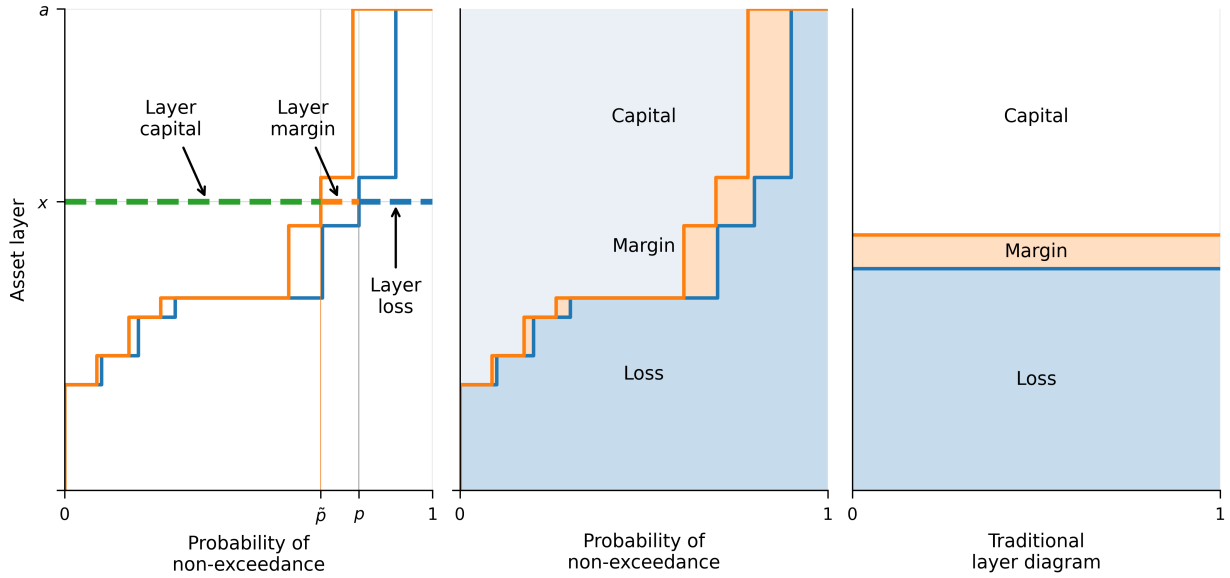


Figure 4.2. Computing loss, margin, and premium for the simple example, CCoC distortion.

In Equation 4.6, the last two formulas can be interpreted as the expected value outcomes X_k with respect to adjusted or **distorted probabilities** q_k , analogous to $p_k = S_{k-1} - S_k$. (This symbol q for a probability measure is not to be confused with capital amount symbol Q .) Notice that $\sum_{j=0}^n q_j = 1$. Since all q_j are nonnegative, collectively they satisfy the properties of a probability distribution.

Exercise. For $E[X]$, prove that $\sum_j S_j \Delta X_j = \sum_j p_j X_j$.

Solution.

$$\begin{aligned} \sum_{j=0}^n p_j X_j &= \sum_{j=1}^n (S_{j-1} - S_j) X_j \\ &= \sum_{j=0}^{n-1} S_j X_{j+1} - \sum_{j=1}^n S_j X_j \\ &= \sum_{j=0}^{n-1} S_j \Delta X_j \end{aligned}$$

because $S_0 X_0 = S_n X_n = 0$. □

Exercise. For $\rho(X)$, prove that $\sum_j g(S_j) \Delta X_j = \sum_j q_j X_j$.

Solution. This is essentially the same derivation as for $E[X]$ but with $g(S)$ substituted for S and q substituted for p . □

We can go one step further and write

$$P = \sum_{k=0}^n q_k X_k = \sum_{k=0}^n X_k Z_k p_k. \quad (4.7)$$

The likelihood ratios $Z_j = q_j/p_j$ individually can be less than, equal to, or greater than one, but their expected value equals one.

Recall that for our example $\iota = 0.15$, $\nu = 0.86956$, and $\delta = 0.13043$. Table 4.4 tabulates the portfolio expected loss and premium based on the CCoC distortion, which we define in Section 4.6. We can see, comparing to Table 4.11, that the same portfolio premium is obtained.

Table 4.4. Pricing the portfolio with CCoC distorted $g(S)$.

k	p	Portfolio X_k	Layer Size ΔX_k	Exceedance Pr S_k	Distorted $g(S)$	$S\Delta X$	$g(S)\Delta X$
0	0	0	22	1	1	22	22
1	0.1	22	6	0.9	0.913	5.4	5.478
2	0.1	28	8	0.8	0.826	6.4	6.609
3	0.1	36	4	0.7	0.739	2.800	2.957
4	0.4	40	15	0.3	0.391	4.500	5.870
5	0.1	55	10	0.2	0.304	2	3.043
6	0.1	65	35	0.1	0.217	3.5	7.609
7	0.1	100	0	0	0	0	0
Sum	1		100			46.6	53.565

4.5. Pricing units via pricing layers

A portfolio's premium, computed by an SRM, has an NA to units.

We are now in a position to consider the pricing of portfolio units. Let the portfolio aggregate loss be the sum of m unit losses:

$$X = \sum_{i=1}^m X^i.$$

The various X^i may or may not be correlated. We assume they are nonnegative.

It is natural to look at Equation 4.6 and suppose that the price of X^i should be

$$P^i = \sum_{j=0}^a q_j X_j^i. \quad (4.8)$$

Note the resemblance of this to the expected value of X^i :

$$L^i = \sum_{j=0}^a p_j X_j^i. \quad (4.9)$$

Finally, of course, unit margin (analogous to Equation 4.1) is simply the difference:

$$M^i = P^i - L^i. \quad (4.10)$$

This decomposition of P into $\sum_i P^i$ is known as the **NA** of an SRM—so-called because it relies on no additional assumptions. It is the **distorted expectation** of X^i . Other ways of writing it are

$$P^i = E[Z \cdot X^i] = E[X^i]E[Z] + \text{COV}(Z, X^i) = E[X^i] + \text{COV}(Z, X^i) \quad (4.11)$$

because $E[Z] = 1$. The first expression resembles a pricing kernel in modern finance theory or a co-measure (Venter et al. 2006). The last expression resembles the capital asset pricing model.

Exercise. Verify Equation 4.11.

Solution. $\text{COV}(Z, X) = E[ZX] - E[Z]E[X] = E[ZX] - E[X]$ because $E[Z] = 1$. \square

Why do we require unique outcome values for X ? To enforce a unique ordering of outcomes. The values q_k depend on the order of events. If there are multiple k with equal X_k , that ambiguity doesn't matter for computing total premium P . But when we consider allocations to units it does, because it allows some unwanted flexibility in q_k . When the outcomes of X are distinct, we are also guaranteed that the NA equals the marginal cost of increasing business in unit i . Since the outcomes are distinct, the order of outcomes of X and of $X + \epsilon X^i$ is the same provided ϵ is small enough (positive or negative). As noted above, q reproduces stand-alone pricing for all risks comonotonic with (same order as) X . Thus

$$\begin{aligned} \nabla^i \rho_X &= \lim_{\epsilon \rightarrow 0} \frac{\rho(X + \epsilon X^i) - \rho(X)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{E_q[X + \epsilon X^i] - E_q[X]}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{E_q[X] + E_q[\epsilon X^i] - E_q[X]}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{\epsilon E_q[X^i]}{\epsilon} \\ &= E_q[X^i]. \end{aligned}$$

This very important result was proved by Delbaen around 2000. It shows that the NA is the *natural* extension of Tasche (1999).

4.6. Five representative distortion functions

We now introduce five carefully selected families of distortion functions: constant cost of capital (CCoC), proportional hazards (PH), Wang transform, dual moment transform, and TVaR. Table 4.5 shows the formulas for each and parameters selected to achieve InsCo's portfolio pricing consistent with an expected return of $\iota = 0.15$. For the Wang distortion, introduced in Wang (2000), Φ is the standard Gaussian cumulative distribution function. For the PH (respectively, other four) a lower (respectively, higher) parameter indicates a more risk averse distortion resulting in higher prices for a given risk.

Table 4.5. Distortion functions calibrated to portfolio price $\rho(X) = 53.565$.

Distortion	Formula	Parameter
CCoC	$\nu s + \delta$	$\iota = 0.1500$
PH	s^α	$\alpha = 0.7205$
Wang	$\Phi(\Phi^{-1}(s) + \lambda)$	$\lambda = 0.3427$
Dual	$1 - (1 - s)^m$	$m = 1.5951$
TVaR	$1 \wedge s/(1 - p)$	$p = 0.2713$

Figure 4.3 plots the five distortion functions and distorted probabilities. The blue star above $s = 0$ on the right reminds us of the probability mass for the CCoC. The thin black line shows the reference $Z = 1$ line, where risk-adjusted and objective probabilities

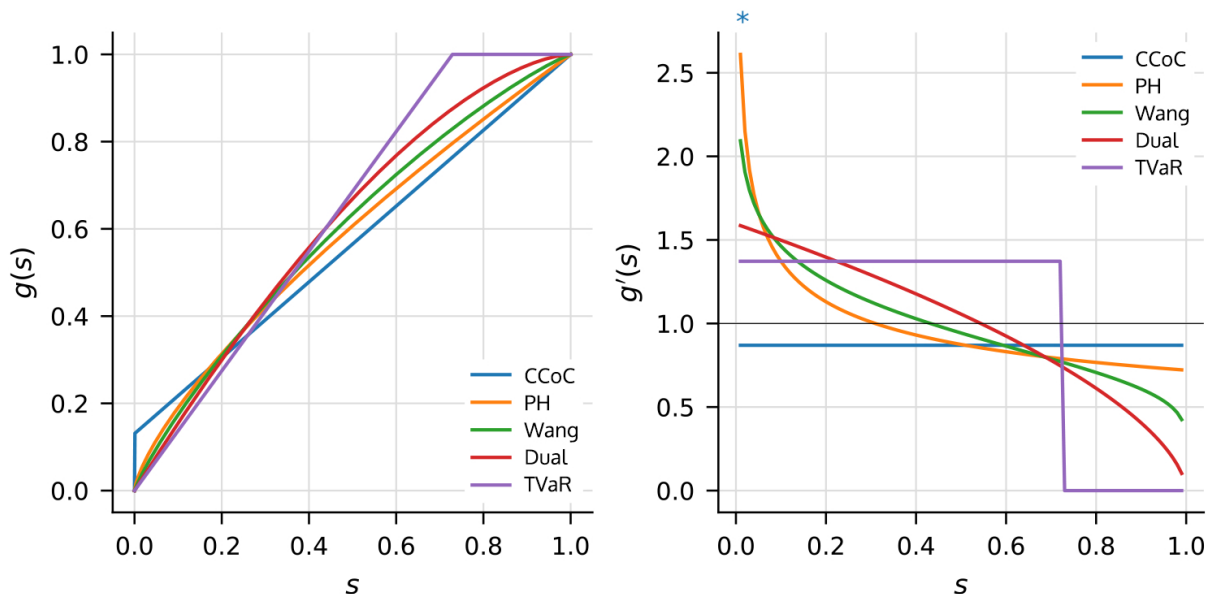


Figure 4.3. Distorted exceedance ($g(s)$, left) and distorted probabilities (q , right) by distortion type.

are equal. The plots explain the order of the distortions: CCoC gives the greatest weight to the worst $s = 0$ loss (remember the * above the worst event) and TVaR the least. Conversely, TVaR gives the greatest weight to moderate losses between the 25th and 75th percentiles. Table 4.6 shows the $g(s)$ values corresponding to each of the five distortions for S values in our example, Table 4.7 shows the distorted probabilities (q) associated with each distortion function, and Table 4.8 shows the $Z = q/p$ factor for each distortion, showing the proportionate increase or decrease over objective probabilities p .

Table 4.6. Distorted exceedance ($g(s)$) by distortion type.

k	S	CCoC	PH	Wang	Dual	TVaR
0	1	1	1	1	1	1
1	0.9	0.9130	0.9269	0.9478	0.9746	1
2	0.8	0.8261	0.8515	0.8819	0.9233	1
3	0.7	0.7391	0.7734	0.8071	0.8535	0.9606
4	0.3	0.3913	0.4200	0.4279	0.4339	0.4117
5	0.2	0.3043	0.3136	0.3089	0.2995	0.2745
6	0.1	0.2174	0.1903	0.1739	0.1547	0.1372
7	0	0	0	0	0	0

Table 4.7. Distorted probabilities (q) by distortion type.

k	p	CCoC	PH	Wang	Dual	TVaR
1	0.1	0.0870	0.0731	0.0522	0.0254	0
2	0.1	0.0870	0.0754	0.0660	0.0513	0
3	0.1	0.0870	0.0781	0.0748	0.0698	0.0394
4	0.4	0.3478	0.3534	0.3791	0.4196	0.5489
5	0.1	0.0870	0.1064	0.1190	0.1344	0.1372
6	0.1	0.0870	0.1233	0.1350	0.1448	0.1372
7	0.1	0.2174	0.1903	0.1739	0.1547	0.1372

Table 4.8. Likelihood ratios ($Z = q/p$) by distortion type.

k	p	CCoC	PH	Wang	Dual	TVaR
1	1	0.8696	0.7310	0.5216	0.2540	0
2	1	0.8696	0.7541	0.6598	0.5134	0
3	1	0.8696	0.7810	0.7480	0.6979	0.3941
4	1	0.8696	0.8834	0.9479	1.0490	1.3723
5	1	0.8696	1.0640	1.1899	1.3439	1.3723
6	1	0.8696	1.2329	1.3502	1.4479	1.3723
7	1	2.1739	1.9033	1.7391	1.5470	1.3723

All of the distortions underweight events 1 through 3. Event 4 has the median loss. TVaR, while assigning zero weight to the lowest losses, is still the most body focused because it overweights events 4 and above—equally—and among all distortions applies the smallest weight to the largest loss. CCoC, however, applies the largest weight to event 7 (the largest loss) and only that event. The other distortions apply weights to event 7 in the order of presentation in the table. The ordering for event 4, representing body focus, is the opposite of the ordering for event 7, representing tail focus.

The results obtained by using the Wang distortion are shown in Table 4.9.

Table 4.9. Pricing the portfolio with Wang transformed probabilities q .

k	p	Unit A X_k^1	Unit B X_k^2	Unit C X_k^3	Portfolio X_k	S	$g(S)$	q	Z
0	0	0	0	0	0	1	1	0	0
1	0.1	15	7	0	22	0.9	0.948	0.052	0.522
2	0.1	15	13	0	28	0.8	0.882	0.066	0.660
3	0.1	5	20	11	36	0.7	0.807	0.075	0.748
4	0.4	10	24	6	40	0.3	0.428	0.379	0.948
5	0.1	26	19	10	55	0.2	0.309	0.119	1.190
6	0.1	17	8	40	65	0.1	0.174	0.135	1.350
7	0.1	16	20	64	100	0	0	0.174	1.739
$L = E_p$		13.4	18.3	14.9	46.6				
$P = E_q$		14.109	18.637	20.818	53.565				
$M = P - L$		0.709	0.337	5.918	6.965				
Prop of tot M		0.102	0.048	0.850	1				

So far, we have seen how to allocate premium P (and therefore margin M , via Equation 4.1) to portfolio units via the NA of an SRM, once a distortion function $g(s)$ is available. Chapter 8 discusses how to choose a distortion function. A premium allocation allows us to compute economic value added by unit, the actual premium in excess of required premium, and to assess static portfolio performance by unit—motivations for performing capital allocation in the first place (Chapter 5). In many ways it is also a good place to stop.

However, if desired, the ideas behind the natural premium allocation can be extended to give a capital allocation consistent with SRM premium allocation. This is discussed in Section 8.3.

4.7. The industry standard approach and its problems

In this section we step back to the 1990s, before the layer approach was formally introduced, and describe how what we call the industry standard approach (ISA) evolved and the problems it generates. At that time, actuaries determined prices on a stand-alone

basis, because in an efficient market prices are additive. They used the formula premium equals loss cost plus margin, and margin was computed as cost of capital times amount of capital, giving (Table 4.1 P1, P3a)

$$P = L + \iota Q = \nu L + \delta a. \quad (4.12)$$

It was deemed obvious that the cost of capital should vary by unit according to different levels of systematic risk. However, attempts to compute underwriting betas from accounting data failed or had standard errors that swamped differences by unit. In consequence, it became standard to assume that ι was constant across units. This story emerges in Myers and Cohn (1987), Cummins (2000), and Kozik (1994).

Around the same time, Tasche (1999) pointed out that marginal capital is the “only definition for the risk contributions which is suitable for performance measurement.” Tasche’s analysis relies on a constant cost of capital to equate *needs more capital* with *needs more margin*. That relationship obviously fails if capital costs vary. As a result, by 2000, the ISA was to use a CCoC and marginal capital by unit to steer portfolios. This is called RAROC, *return on risk-adjusted capital*: the return is constant but the capital varies with risk.

As we saw in Section 4.4, in the mid-1990s Wang and others introduced the idea of pricing by layer. It is easy to translate Equation 4.12 into the layer framework. Without loss of generality, we can assume the four adjustments of Section 4.3, so $a = \max(X)$. Then,

$$\begin{aligned} P &= L + \iota Q \\ &= \nu L + \delta a \\ &= \nu E[X] + \delta \max(X) \\ &= \nu \sum_{k=0}^{n-1} S_k \Delta X_k + \delta \sum_{k=0}^{n-1} \Delta X_k \\ &= \sum_{k=0}^{n-1} (\nu S_k + \delta) \Delta X_k \\ &= \sum_{k=0}^{n-1} g(S_k) \Delta X_k \end{aligned}$$

is exactly the premium associated with the CCoC distortion $g(s) = \nu s + \delta$. Here an important transformation has occurred. ISA (reasonably) assumed a CCoC *by unit*. The layer view shows that constant unit costs mathematically imply a CCoC *by layer*, an implication that went unremarked at the time. The problem is that **capital costs are manifestly not constant by layer**. Numerous websites such as the Federal Reserve’s FRED and Bank of America’s ICE provide the latest information on credit yield spreads. Low-



Figure 4.4. Assuming a constant cost of capital. © S. Mildenhall, 2024.

risk instruments such as government or AAA-rated bonds require lower yields than non-investment-grade or junk bonds. This is true even after corrections are made to the yield calculations to take default probability into account (Heynderickx et al. 2016).

Aside from the factual inconvenience of making an empirically false assumption, the CCoC distortion causes other problems. It is numerically unstable because it puts a large weight on the worst outcome and underweights all others. For the same reason, it is very averse to tail risk but blithely ignores volatility risk around the mean. None of the four other distortions in Table 4.5 has these problems: all have a cost of capital that varies by layer and distribute their probability weight more evenly across events. CCoC is further unusual because it provides no diversification benefit for independent risks.

Exercise. Show that CCoC distortion pricing is additive for independent risks.

Solution. $P = \nu E[X] + \delta \max(X)$ and expectation and maximum are additive for independent risks. \square

Table 4.10 shows the calculation of q_k , $E[X]$, and the price of X using the q s. It gives the same result as Table 4.4.

Let us examine the allocations obtained by the CCoC assumption and compare the results obtained by using the Wang distortion.

The Z values associated with the Wang distortion are shown in Table 4.9; the CCoC values are in Table 4.10. The last three Z values are greater than one for Wang, whereas only the last one is greater than one for CCoC.

Table 4.10. Pricing the portfolio with CCoC distorted probabilities q .

k	p	Portfolio X_k	Layer Size ΔX_k	Exceedance Pr S_k	pX	Distorted q	qX	$Z = q/p$
0	0	0	22	1	0	0	0	
1	0.1	22	6	0.9	2.2	0.087	1.913	0.870
2	0.1	28	8	0.8	2.800	0.087	2.435	0.870
3	0.1	36	4	0.7	3.6	0.087	3.130	0.870
4	0.4	40	15	0.3	16	0.348	13.913	0.870
5	0.1	55	10	0.2	5.5	0.087	4.783	0.870
6	0.1	65	35	0.1	6.5	0.087	5.652	0.870
7	0.1	100	0	0	10	0.217	21.739	2.174
Sum	1		100				53.565	

For InsCo, the results of the CCoC allocation are shown in Table 4.11. Unit C, because of its outsize presence in the worst event, gets 91.9% of the capital and margin allocated to it. This tends to be the outcome from practical applications of the standard approach: the thick-tail lines of business get nearly all the capital and margin allocated to them and the thin tail lines get a free pass. With the Wang distortion, instead of getting 8% of the margin, Units A and B get 15% of the margin.

Table 4.11. Industry standard approach to allocation.

Variable	Unit A X_j^1	Unit B X_j^2	Unit C X_j^3	Total X_j
$L = E[X]$	13.4	18.3	14.9	46.6
$a = \text{TVaR}$	16	20	64	100
$Q = \nu\text{XTVaR}$	2.261	1.478	42.696	46.435
P	13.739	18.522	21.304	53.565
$M = P - L$	0.339	0.222	6.404	6.965
Allocation of M	0.049	0.032	0.919	1
$\iota = M/Q$	0.150	0.150	0.150	0.150

To recap: It is often assumed that one can transition from allocating capital to allocating cost of capital (i.e., margin or premium) simply by multiplying capital by a single cost of capital rate. Indeed, it *is* basically that simple when dealing with the entire business portfolio in aggregate (because the total cost of capital is the weighted average cost times the amount). However, when disaggregating results into units, this simple tactic founders for two reasons. First, as we saw above, purported capital allocation procedures based on marginal analysis of the capital risk measure aren't really relevant to pricing. Second, it is likely that CCoC is not appropriate because there is no single cost of capital rate that applies to all layers of risk and different units consume a different mix of cap-

ital by layer. Each contributor to the business has its own risk profile and its own joint dependency with the aggregate risk profile—and therefore its own cost of capital.

Understanding the flaws of the ISA highlights the improvements offered by layer pricing and allocation. First, by choice of distortion g , we can match the cost of capital by layer to market observables. Second, since unit consumption of capital varies by layer, the calculation naturally produces different returns by unit in total, thus solving the problem from the 1990s originally tackled using underwriting betas. Moreover, the solution is consistent with the observations in Table 1.1. Thick-tailed, catastrophe exposed lines consume relatively more higher layer capital. Higher layer capital is less exposed to loss (like a higher rated bond) and earns a lower return. Thus, such lines will be charged with a lower cost of capital. They will run with lower leverage too, resulting in a low loss ratio. This is why catastrophe reinsurance works economically: it substitutes expensive equity for cheaper debt-like capital.

4.8. Pricing events: An explanation

We saw in Equation 4.6 that premium can be computed as the expected value of the loss outcomes X_k with respect to altered probabilities q_k . This seems to provide us with an interpretation of pricing events, rather than layers, but that interpretation comes with a big caveat.

So far, this is just algebra. We would *like* to interpret Z_k as the unique **state price** for event k and take this as the definitive statement of event pricing, backed by a no-arbitrage argument. Alas, that interpretation is not quite correct for reasons that involve bid-ask spreads and are beyond the scope of this monograph. They are explained more fully in Chapter 10 of Mildenhall and Major (2022).

However, it does provide a set of internally consistent allocations to events as part of the total. The portfolio premium P can be regarded as an expectation taken under the new q probability $E_q[X]$ or, equivalently, as a *weighted* expectation $E_p[ZX]$ under the original distribution with Z as the weight. Using the q probabilities this way is also applying an SRM to X , equivalent to how we did it previously through layer pricing. Using a distorted expectation to represent price is very much in the spirit of using risk neutral probabilities in modern finance; they are not always unique there, either. Thus, if we fix X , then we can price any subunit X^i of X using q_k

$$\text{allocation to } X^i = \sum_k X_k^i q_k$$

and never run into logical or economic inconsistencies. This is *exactly the same* as the procedure for computing coTVaR and is the basis of the NA.

In general, Equation 4.7 can also be used to price any set of outcomes Y_k that are comonotonic with X_k —in the sense that it will give the same stand-alone price as the spectral measure. This follows because the ordering of events determined by X_k gives a nondecreasing ordering of events for Y_k . (The converse may not be true. For example, a constant is comonotonic with any X_k , but would admit any ordering of events as nondecreasing.) In our case, each layer is comonotonic with the total, so Equation 4.7 can be seen as an application. In general, the allocation will be less than or equal to the stand-alone price. There is a lower bound too; see Chapter 10 of Mildenhall and Major (2022).

4.9. Summary

Spectral pricing gives the price of the total (whole portfolio) risk in the market. In contrast, the NA gives an allocation to part of the total risk—and crucially depends on the total. The allocation equals the stand-alone spectral price only in the case that the risk is comonotonic with the total risk. You can use and interpret the NA as a risk-adjusted expected value, and never have a problem, provided you understand it in the context of a total X . As an allocation, it has the beauty of being consistent with a marginal approach in the case that all the outcomes of X are distinct. In fact, it is consistent if and only if all the outcomes are distinct.

Spectral prices are subadditive, so insureds benefit from pooling together—providing a motivation for insurance companies to exist! The insurer has to figure out how to allocate that lower price fairly among its insureds. Of course, the insurer is happy with any price greater than or equal to a fair allocation, particularly with the marginal interpretation of allocation. The marginal interpretation drives profit maximizing behavior in a standard sense.

We conceptualized the cash flows between policyholders and investors via the InsCo one-year model with claim payouts X and probabilities p . We distinguished between capital and pricing risk measures. Portfolio funding was subject to the constraints $P+Q = a$ and $P = L + M$ with the investor return being $\iota = M/Q$, resulting in Equation 4.2:

$$P = L + \delta(a - L) = L + \iota Q.$$

These relationships were carried down to hold for each asset layer. Pricing was done via SRMs through distorted probabilities with S being replaced by $g(S)$ and p being replaced by $q = \Delta g(S)$. This allows us to price any cash flow that is a function of specific value outcomes X . In particular we can price the units of the portfolio: lines of business or even policies.

We saw that while understanding the marginal impact of increasing a unit's exposure on the overall asset requirement (the capital measure) was meaningful for risk manage-

ment, it had little relevance to pricing *per se*. We also saw that assuming a CCoC had troublesome implications about event weights $Z = q/p$.

4.10. Algorithms

The next two algorithms are taken from Mildenhall and Major (2022). They are included here for completeness and ease of reference. As is often the case, they look more intimidating than they are—build the spreadsheet and you’ll see they are quite straightforward!

4.10.1. Algorithm to evaluate an SRM on a discrete random variable

Algorithm Input: X is a discrete random variable, taking values $X_j \geq 0$, and $p_j = P(X = X_j)$, $j = 1, \dots, n$. ρ_g is an SRM.

Follow these steps to determine $\rho_g(X)$.

Algorithm Steps

- (1) **Pad** the input by adding a zero outcome $X_0 = 0$ with probability 0.
- (2) **Sort** events by outcome X_j into ascending order.
- (3) **Group** by X_j and sum the corresponding p_j . Relabel events $X_0 < X_1 < \dots < X_{n'}$ and probabilities $p_0, \dots, p_{n'}$. All X_j are distinct.
- (4) **Decumulate** probabilities to determine the survival function $S_j := S(X_j)$ using $S_0 = 1 - p_0$ and $S_j = S_{j-1} - p_j$, $j > 0$.
- (5) **Distort** the survival function, computing $g(S_j)$.
- (6) **Difference** $g(S_j)$ to compute risk-adjusted probabilities $\Delta g(S_0) = 1 - g(S_0)$, $\Delta g(S_j) = g(S_{j-1}) - g(S_j)$, $j > 0$.
- (7) **Sum-product** to compute $\rho_g(X) = \sum_j X_j \Delta g(S_j)$.

Steps (4) and (6) are inverse to one another. The transition $p \rightarrow S \rightarrow gS \rightarrow \Delta gS$ from input probabilities to risk-adjusted probabilities is used in all SRM-related algorithms.

The algorithm computes the outcome-probability formula

$$\rho_g(X) = \int_0^\infty x g'(S(x)) dF(x) = \sum_{j>0} X_j \Delta g(S_j).$$

When X is discrete the Steiltjes integral becomes a sum—the joy of using dF . We can also use the survival function form with $\Delta X_j = X_{j+1} - X_j$:

$$\rho_g(X) = \int_0^\infty g(S(x)) dx = \sum_{j \geq 0} g(S_j) \Delta X_j.$$

4.10.2. Algorithm to compute the linear NA for discrete random variables

Algorithm Inputs:

- (i) The outcome values (X_j^1, \dots, X_j^m) , $j = 1, \dots, n$, of a discrete m -dimensional multivariate loss random variable. Outcome j occurs with probability p_j . $X_j = \sum_i X_j^i$ denotes the total loss for outcome j .
- (ii) An SRM ρ_g associated with the distortion function g .

Follow these steps to determine $D^n \rho_X(X^i)$, the NA of $\rho(X)$ to unit i .

Algorithm Steps

- (1) **Pad** the input by adding a zero outcome $X_0^1 = \dots = X_0^m = X_0 = 0$ with probability $p_0 = 0$.
- (2) **Sort** events by total outcome X_j into ascending order.
- (3) **Group** by X_j and take p -weighted averages of the X_j^i within each i and $X_j = x$ group. Sum the corresponding p_j . Relabel events using $j = 0, 1, \dots, n'$ as X_j^i and probabilities $p_0, \dots, p_{n'}$.
- (4) **Decumulate** probabilities to determine the survival function $S_j := S(X_j)$ using $S_0 = 1 - p_0$ and $S_j = S_{j-1} - p_j$, $j > 0$.
- (5) **Distort** the survival function, computing $g(S_j)$.
- (6) **Difference** $g(S_j)$ to compute risk-adjusted probabilities $\Delta g(S_0) = 1 - g(S_0)$ and $\Delta g(S_j) = g(S_{j-1}) - g(S_j)$, $j > 0$.
- (7) **Sum-product** to compute $\rho_g(X) = \sum_j X_j \Delta g(S_j)$ and

$$D^n \rho_X(X_i) = E[X_i Z] = \sum_j X_j^i Z_j p_j = \sum_j X_j^i \frac{\Delta g(S_j)}{p_j} p_j = \sum_j X_j^i \Delta g(S_j). \quad (4.13)$$

Comments

- (a) When the data are produced by a simulation model, n equals the number of events and m the number of units. With realistic data, n is in the range thousands to millions, and m ranges from a handful up to the hundreds for a full corporate model.
- (b) Step (1) only results in a new outcome row when the smallest X_j observation is > 0 .
- (c) The averages in Step (3) are implemented as

$$\kappa^i(x) = E[X_i | X = x] = \frac{\sum_{j: X_j = x} p_j X_j^i}{\sum_{j: X_j = x} p_j}.$$

- (d) After Step (3), the X_j are distinct, they are in ascending order, $X_0 = 0$, and $p_j = P(X = X_j)$.
- (e) The backward difference $\Delta g(S_j)$ computed in Step (6) replaces $g'(S)dF(x)$ in various formulas.
- (f) Equation 4.13 is an exact equation for a discrete distribution. Approximation occurs if it is applied to the empirical distribution of a discrete sample representing a different underlying distribution, possibly one with density.

5. Applications of pricing and allocation

Performance assessment, new business pricing, and reinsurance decision-making are all essentially the same exercise as price allocation. Mergers and acquisitions, and portfolio optimization, must be approached differently.

Table 5.1 summarizes the events affecting InsCo's three units of business as well as an available 35 excess of 65 stop loss reinsurance cover. This has ceded losses of zero in events 0 through 6 and 35 in event 7. It has an expected payout of 3.5 (loss on line 10%) and ceded premium of 6.5 (rate on line 18.6%). The expected net profit to the reinsurer is 3.0.

In the last three rows of the table we see the expected loss, the plan premium, and the required premium as calculated by a Wang transform calibrated to achieve a 15% return on capital for the (gross) portfolio.

Table 5.1. Events with unit losses and statistics.

k	p	Unit A	Unit B	Unit C	Gross	35 xs 65	Net
0	0.0	0	0	0	0	0	0
1	0.1	15	7	0	22	0	22
2	0.1	15	13	0	28	0	28
3	0.1	5	20	11	36	0	36
4	0.4	10	24	6	40	0	40
5	0.1	26	19	10	55	0	55
6	0.1	17	8	40	65	0	65
7	0.1	16	20	64	100	35	65
$E[X]$		13.400	18.300	14.900	46.600	3.500	43.100
Plan		13.900	18.700	19.600	52.200	6.500	45.700
ρ_{Wang}		14.109	18.637	20.819	53.565	6.087	47.478

5.1. Performance assessment

Which units are performing at a satisfactory level?

Performance assessment is a multiattribute question, involving:

- Recent and prospective growth
- Expense control

- Employee turnover
- Customer retention
- Competitive standing
- Product and operational innovation

and other qualities. However, a key factor is profitability. The first question is always: is this unit meeting a profit benchmark?

Recall we have a 15% return benchmark for the (gross) portfolio as a whole. Table 5.1 shows that additional premium of $53.565 - 52.2 = 1.365$ is required to meet this threshold. This is true independent of the choice of distortion function—it is a direct consequence of the 15% return on capital requirement. Here we assume a calibrated Wang transform is used to allocate the portfolio requirement down to the units.

Table 5.2 shows that of the three units, only Unit B has a plan premium that exceeds the allocated target premium, and there, only by 0.063. Unit A has a deficit of 0.209, and Unit C has most of the portfolio deficit at 1.219. Putting these shortfalls in perspective, Unit C still has the highest gap measured in loss ratio (4.45%), but Unit A needs to increase its plan margin by the greatest factor (41.8%).

Table 5.2. Plan premium relative to target required premium.

Metric	Unit A	Unit B	Unit C	Gross
Plan premium	13.900	18.700	19.600	52.200
ρ^{Wang}	14.109	18.637	20.819	53.565
$E[X]$	13.400	18.300	14.900	46.600
Plan margin	0.500	0.400	4.700	5.600
Excess margin	-0.209	0.063	-1.219	-1.365
% of plan	-41.8%	15.6%	-25.9%	-24.4%
Excess LR	-1.43%	0.33%	-4.45%	-2.28%

Note this analysis depends heavily on the fact that we have chosen the Wang transform to drive our SRM allocation. What if a different distortion function were used? And how do we decide which one to use? These questions are addressed in Chapter 8.

5.2. Reinsurance decision-making

Reinsurance serves as loss protection and a source of loss-bearing capital. Table 5.1 shows a 35 excess of 65 portfolio stop loss contract. We assume this is aggregate cover, so questions of reinstatement do not apply. Per-occurrence cover would need to be analyzed to see its impact on aggregate results.

Should this cover be purchased?

Reinsurance decisions must weigh several factors, including impact on capital requirements and underwriting constraints, counterparty credit risk, and price. First let us consider capital requirements.

Previously, we assumed a $\text{TVaR}_{0.99}$ standard for required assets. For the gross portfolio, this led to $a = 100$, full collateral in this example. For the net portfolio, it would lead to $a = 65$ if full credit were to be allowed or if the same standard were applied to net losses.

Exercise. Apply the rate of return-driven pricing formula (Equation 4.2) to derive the required premium for the net portfolio based on a 15% required return. What might you conclude about the price of the reinsurance cover?

Solution.

$$P = \nu L + \delta a = (0.86956)(43.1) + (0.13043)(65) = 45.956,$$

where $L = 43.1$ is the net expected loss per Table 5.1, and $a = 65$ is the required asset conditional on having stop loss reinsurance. For comparison, recall from the end of Section 4.1 that $P(\text{gross}) = 0.86956 \cdot 46.6 + 0.13043 \cdot 100 = 53.565$ (consistent with plugging in values from the Gross column of Table 5.1). This represents savings of 7.609, which is the maximum that InsCo should be willing to cede for the cover. Since the cover costs 6.5 (Table 5.1), we might conclude that the price of the cover is advantageously low and it should be purchased (unless a better deal were available). \square

Unfortunately, this line of thinking relies on the familiar CCoC assumption. The risk profile has changed between the gross and net portfolio, so there is no basis for assuming that the same rate of return is required. An alternative approach is to use the loss protection interpretation and view reinsurance as a source of capital. This is elaborated upon in Mildenhall and Major (2022).

Applying the same Wang transformation allocation to the reinsurance cash flows, Table 5.1 shows a target required premium of 6.087, which is 0.413 *lower* than the actual 6.5 ceded premium. We conclude the cover should *not* be purchased.

Exercise. Using the Wang allocation and assuming $a = 65$, compute the return on capital for the net portfolio. Explain why the result is in this direction.

Solution.

$$\iota = \frac{M}{Q} = \frac{P - L}{a - P} = \frac{47.478 - 43.1}{65 - 47.478} = \frac{4.378}{17.522} = 25\%.$$

All investment tranches $(Q_j^{\text{Net}}, M_j^{\text{Net}}) = (Q_j^{\text{Gross}}, M_j^{\text{Gross}})$, $j = 1, \dots, 5$, are identical, but the last tranche $j = 6$ has $\Delta X_j^{\text{Net}} = 0 < \Delta X_j^{\text{Gross}}$. Relative to gross, the net weighted average cost of capital is therefore skewed to the lower tranches, which have higher exceedance probabilities and therefore higher required returns. \square

5.3. New business pricing

While underwriting involves many factors other than pricing, such as exposure guidelines and physical and financial standards reflecting moral hazard, etc., pricing is a key task in the onboarding of new business.

At first glance, this seems equivalent to the performance assessment task: calculating a required premium as a distorted loss expectation. It would seem there is not much to add to this discussion. Typically, however, proposed new business is not run through the full capital model directly because that would be inefficient in terms of computing resources and timeliness. Rather, a simplified or surrogate version of the capital model is used to extract key loss statistics.

Let's say for example we have a (relatively) simple cat model through which the portfolio has been run—once, and the results stored—and through which a new business prospect has just been run. Non-cat losses are represented by an independent lognormal variable calibrated on exposure characteristics. The resulting scatter plot is shown in Figure 5.1.

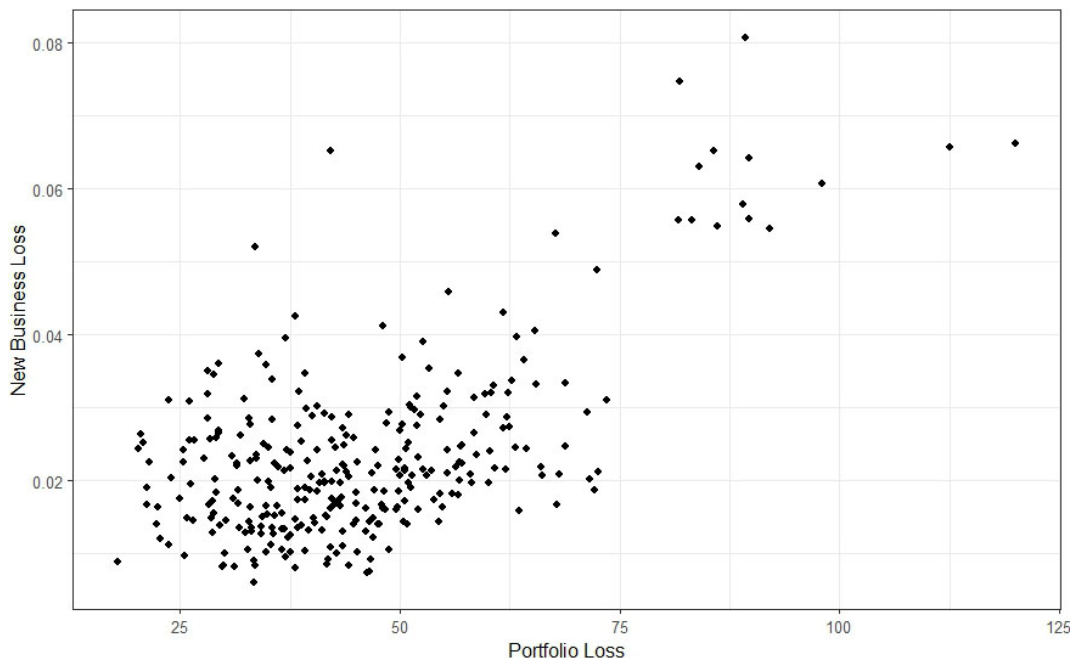


Figure 5.1. Simulated losses from a surrogate model applied to the gross portfolio and a new business prospect.

The problem is to translate these results into a form suitable for SRM allocation pricing. One could use the calibrated Wang transform to sort these results by the portfolio loss and calculate distorted probabilities, but we have no guarantee that the surrogate model accurately reproduces the distribution of portfolio losses that the capital model uses. The solution is to extract the conditional expected new business losses for each portfolio loss in Table 5.1. (See the discussion of $\kappa^i(X_k)$ in Section 4.3).

This extraction can be done via local regression (LOWESS, LOESS, kernel smoothing; Friedman et al. (2008)). Figure 5.2 shows the scatter plot of Figure 5.1 with an estimate of the conditional mean loss curve and the specific event points. This leads to the augmented loss (Table 5.3).

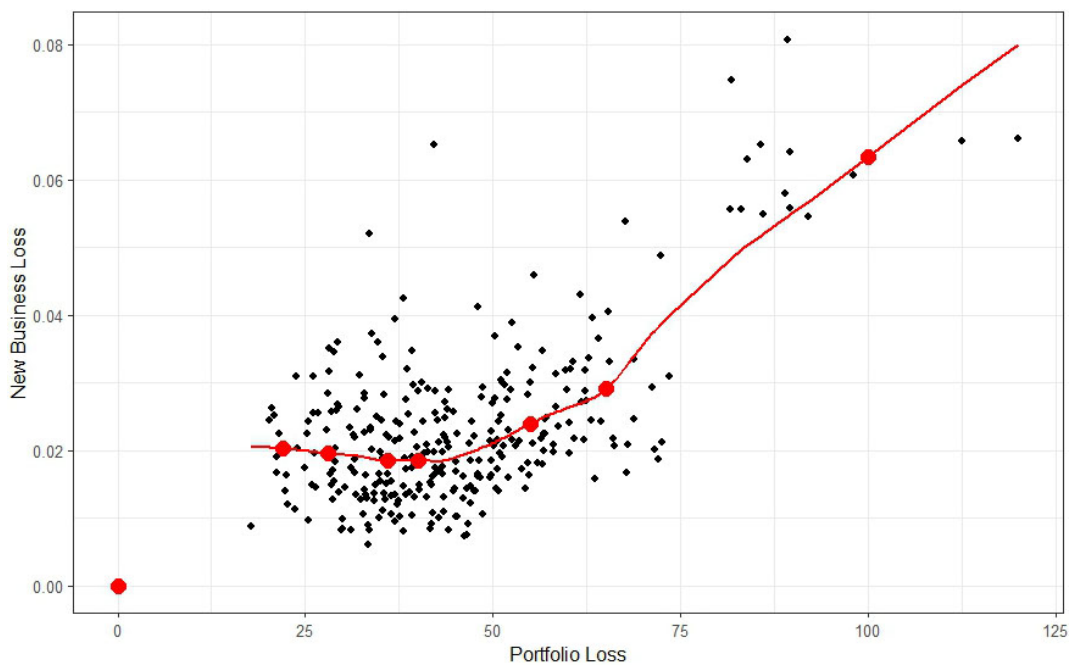


Figure 5.2. Estimating conditional expected losses using locally weighted regression (LOWESS).

Exercise. Why is it reasonable to treat a small new business addition as if it were a unit within the portfolio, using the SRM allocation to calculate a reservation price?

Solution. Because it would have little or no impact on the sort order of portfolio losses. The distorted event probabilities q are functions of the distorted exceedance probabilities $g(S)$, which are in turn functions of exceedance S , which in turn is crucially dependent on the sort order of portfolio event losses. In the simple example, adding the new business losses X^n to portfolio losses X did not change the sort order of $X + X^n$ compared to X . In general, with a realistically fine-grained simulation, there may indeed be changes in the sort order, but they will be small and not be material to the results. \square

Table 5.3. Events with new business conditional expected losses.

k	p	q	Gross	New Business
0	0.0	0.0000	0	0.00000
1	0.1	0.0522	22	0.02028
2	0.1	0.0660	28	0.01965
3	0.1	0.0748	36	0.01849
4	0.4	0.3791	40	0.01855
5	0.1	0.1190	55	0.02402
6	0.1	0.1350	65	0.02910
7	0.1	0.1739	100	0.06339
$L = E[X]$			46.600	0.02491
$P = \rho_{\text{Wang}}(X)$			53.565	0.02858
$M = P - L$			6.965	0.00367

5.4. Mergers and acquisitions

What do you need to consider in a merger or acquisition? In general, there are quite a few considerations, such as strategic fit, financial performance of the target company, cultural compatibility, synergies and cost savings, regulatory compliance, tax implications, valuation for the target company, and risk management.

While the finance department will no doubt have an opinion on valuation, the capital modeling team could also be tasked with an analysis of how the risk profile of the new portfolio fits with the existing portfolio. The same principles discussed previously still apply, but with a twist. Whereas it was reasonable to treat a new business policy as if it were already a (relatively small) unit within the portfolio, a sizable acquisition requires a different approach.

Say instead of the loss distribution portrayed in Figure 5.1, we had an acquisition opportunity with losses scaled up nearly 2,000 times larger on the y -axis at every point, having a mean loss of 48.35. We will focus on losses, with the acquired premiums to be compared to the required premium we will calculate.

Rather than the conditional mean loss of the acquired business (hereinafter X^a), we need the entire conditional distribution (or at least a representative sample). There are several technical approaches to this, such as quantile regression (Koenker and Hallock 2001) or kernel smoothing (Wand and Jones 1994). For our example, we will take what is possibly the simplest nontrivial approach.

1. Calculate the conditional mean $\bar{X}^a = E[X^a|X]$ via LOWESS at every X point in the scaled-up equivalent to Figure 5.1, just as we did to obtain Figure 5.2.
2. At each such X point, compute the squared difference between X^a and \bar{X}^a .

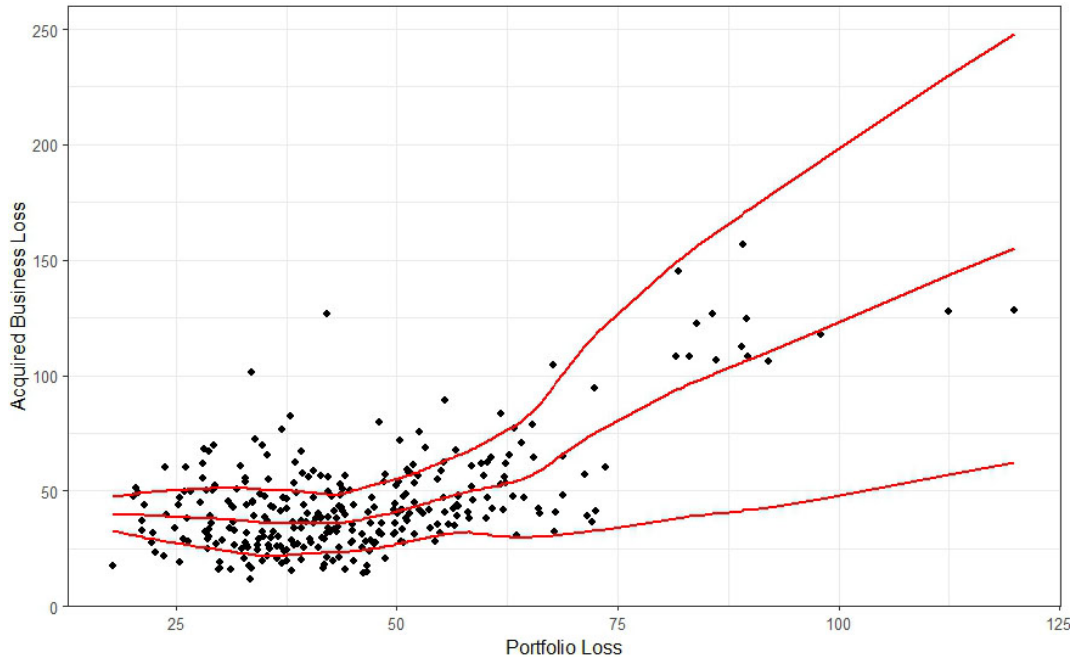


Figure 5.3. Representing the conditional loss distribution using locally weighted regression (LOWESS): mean plus or minus one standard deviation.

3. Use LOWESS again to compute the conditional mean square deviation, i.e., the variance, and its square root, the conditional standard deviation, σ_a .
4. Inspired by Gauss-Hermite quadrature (Kovvali 2022), we compute \bar{X}^a plus or minus one standard deviation. This is depicted in Figure 5.3.
5. Duplicate each row in Table 5.1, inserting a new column for loss X^a consisting of those $\bar{X}^a \pm \sigma_a$ points. Rescale each probability by half.
6. Insert a new column for merger losses $X^m = X + X^a$. Sort on X^m .
7. Compute S , $g(S)$, and q as before, based on the same Wang transform SRM. This is shown in Table 5.4.

The required premium split between our original portfolio X (52.74) and the acquired X^a (59.09) is irrelevant to this analysis! Recall the original portfolio required premium was 53.57. What is important is the difference between the new merger portfolio premium 111.83 and the original premium of 53.57, i.e., 58.26. This is what InsCo would require for premium from the new business to secure enough margin to satisfy investors.

Exercise. Explain the capital situation. What increment of assets is needed with the newly acquired business? What capital? What is the total return on capital for the combined business? Return on incremental capital? Explain why this is not the 15% return originally specified.

Table 5.4. Events with acquired business Gauss-Hermite two-point sampled conditional losses.

k	p	X	X^a	X^m	S	$g(S)$	q
0	0	0	0	0	1	1	0
1	0.05	22	29.37	51.37	0.95	0.9766	0.0234
2	0.05	28	25.47	53.47	0.9	0.9478	0.0287
3	0.05	36	21.55	57.55	0.85	0.9161	0.0318
4	0.2	40	22.74	62.74	0.65	0.7667	0.1494
5	0.05	22	49.35	71.35	0.6	0.7244	0.0423
6	0.05	28	50.79	78.79	0.55	0.6802	0.0442
7	0.05	55	30.83	85.83	0.5	0.6341	0.0461
8	0.05	36	50.22	86.22	0.45	0.5859	0.0482
9	0.2	40	49.27	89.27	0.25	0.3700	0.2159
10	0.05	65	29.89	94.89	0.2	0.3089	0.0611
11	0.05	55	62.38	117.38	0.15	0.2439	0.0650
12	0.05	100	47.73	147.73	0.1	0.1739	0.0700
13	0.05	65	83.06	148.06	0.05	0.0964	0.0775
14	0.05	100	198.30	298.30	0	0	0.0964
$L = E[X]$		46.60	48.35	94.95			
$P = \rho_{\text{Wang}}(X)$		52.74	59.09	111.83			

Solution. Using a 95% TVaR criterion, total assets of 298.30 are needed for the combined portfolio, which is an increase of 198.30 over the original 100. With required premium of 111.83, this corresponds to capital of 186.46 for the combined portfolio, an increase of 140.03 over the original 46.43 ($= 100 - 53.57$) capital. Margin on the combined portfolio is 16.88, for a return of 9.06%. The margin has gone up by 9.92, corresponding to a return of 7.08% on incremental capital. More capital is backing more remote probability losses, requiring less margin and therefore lower returns on capital. \square

5.5. Portfolio optimization

*If this is the best of possible worlds,
what then are the others?
—Voltaire, “Candide”*

5.5.1. Theory

Portfolio management is an inverse problem to pricing. Rather than taking the portfolio as given and determining appropriate pricing, the problem becomes to take the pricing as given and determine the best (or at least a better) composition of the portfolio. This task is typically complicated, however, by constraints such as volume or risk measures.

The textbook approach to optimization (e.g., Postek et al. (2024b), available online at Postek et al. (2024a)) starts with an objective function $f(x)$ where x is typically a vector parameter. The problem is to find the value of x that maximizes (or minimizes) $f(x)$. This is often complicated by the existence of multiple constraint functions $h_j(x)$ requiring $h_j(x) \leq 0$ for $j = 1, \dots, n$ and $h_j(x) = 0$ for $j = n + 1, \dots, m$.

What should be our objective function? It is tempting to think it should be margin, the difference between plan premium and expected loss. However, if our revelations in Chapter 4 taught us anything, it is that modifying the portfolio will modify the risk and that being properly rewarded for risk goes beyond expected loss. We need to compare the (rescaled) plan premium to the (rescaled) required premium from the pricing functional.

We found at the end of Section 4.1 that InsCo's economic value added (EVA) is -1.4 . Our goal will be to make EVA as large (positive) as possible, or at the very least make it less negative.

For portfolio optimization, the particulars will likely look like:

- Objective function is economic value added.
- The parameters are scale factors, one for each unit.
- The primary constraint is to stay at the current capital.
- Secondary constraints are limits on growing or shrinking each unit.

Scale factors determine the volume of each unit. The current portfolio includes the random variable X^i representing the losses from unit i . We now introduce the random variable $X^i(\epsilon)$ to represent its losses if the unit volume has changed by the factor $1 + \epsilon$. The original portfolio loss variable is $X = \sum_i X^i(0)$ in this notation. If unit i were to grow, say, 5%, its loss random variable would be represented as $X^i(0.05)$.

Let the portfolio loss random variable be $X(\epsilon^1, \dots) = \sum_i X^i(\epsilon^i)$ and denote the corresponding required assets by $a(\epsilon^1, \dots)$. For any particular ϵ values, the capital model can compute these.

For the InsCo example, we will assume the constraints are $-0.5 \leq \epsilon \leq 0.5$.

If unit losses scaled linearly, as they do in stock portfolios, then we would have $X^i(\epsilon) = (1 + \epsilon)X^i(0)$. Unfortunately, this is usually not the case with insurance, although catastrophe cover in a limited geographic area comes close. However, a properly defined scale factor will give us scaling in planned premium.

There is a second source of nonlinearity here. As units grow or shrink, so do their losses in each event. Because the computation of the pricing SRM (and probably the capital measure as well) depends on the ordering of portfolio losses, that ordering may change, and so the measure will vary nonlinearly with the scale factors. For small changes in scale (i.e., ϵ close to 0), this reordering may not happen or may not make a material difference. Our approach *will* take this into account.

The primary capital constraint listed above needs elaboration. As the portfolio composition shifts, the asset requirement (specified by the capital risk measure $a(X)$) will change. So will the total plan premium. We interpret the constraint to mean that the portion of assets not supplied by premium, $Q = a - P$ (Equation 1.1), be maintained at the current level. The constraint might have been simply not to exceed the current level, but we don't want to "leave money on the table" in the sense of holding more capital than we are currently using to support risk. So the constraint is an equality. In some other application, there could be a target capital different than the current level. Our approach is easily adaptable to that.

In a real application, there might be other constraints. For example, we might not want total portfolio margin to go down. We will keep the example here relatively simple.

If EVA and required assets scaled linearly, then the above would be a linear programming (LP) problem that could be solved with readily available linear programming codes. Locally, that is to say with ϵ close to zero, this should be a good approximation. However, we want to handle bigger steps correctly, so we need a nonlinear optimization strategy. The Nelder-Mead Simplex algorithm is generally applicable, but its performance may suffer in comparison to methods that use first derivatives such as the conjugate gradient algorithm by Polak and Ribière and the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno. Other methods, such as the Newton conjugate gradient algorithm, use both first and second derivatives.

In Python, these and other methods are available in `scipy.optimize`. In R, the same holds for `nloptr`. Some of these implementations rely on FORTRAN codes written half a century ago!

If we treat losses as scaling linearly, then we can make use of first derivatives in our solution method. First derivatives are almost automatically provided when the objective—portfolio EVA—is evaluated. Second derivatives are impractical to compute here. If we do not wish to treat losses as scaling linearly, then we have a few options:

1. Use a method like Nelder-Mead that does not require derivatives.
2. Compute derivatives numerically by evaluating the objective at nearby points. This could be computationally inefficient.
3. Use a **proxy model**—a simpler version of the complete model that approximates it—to provide derivatives. This is essentially what we are doing when we assume linear scaling, but a more accurate nonlinear proxy might be available.

The overall solution strategy here is stepwise hill climbing:

1. Identify promising directions for scale changes.
2. Select new scale parameters satisfying the capital and scale constraints.
3. Rerun the model at the new scales.
4. Repeat until EVA improvement seems impossible.

With linear scaling, the important thing to remember is that allocated quantities (premium, assets, capital) are the same as marginal quantities. If a unit is scaled up by 1% ($\epsilon = 0.01$), say, then the allocated quantity will go up by 1%.

5.5.2. *InsCo optimization*

The input data for InsCo is provided in Table 5.5.

Table 5.5. InsCo marginal EVA and capital.

Quantity	Unit A	Unit B	Unit C	Portfolio
Plan premium P_p	13.9000	18.7000	19.6000	52.2000
Required premium P_R	14.1092	18.6375	20.8186	53.5652
EVA $V = P_p - P_R$	-0.2092	0.0625	-1.2186	-1.3652
Required assets a	16.0000	20.0000	64.0000	100.0000
Plan capital $Q_p = a - P_p$	2.1000	1.3000	44.4000	47.8000
EVA/capital ratio $r = V/Q_p$	-0.0996	0.0481	-0.0274	
Minimum scale ϵ_-	-0.5	-0.5	-0.5	
Maximum scale ϵ_+	0.5	0.5	0.5	

Exercise. Assume that event 10 is always the one to produce the largest portfolio loss and therefore the capital measure. Compute an expression for ϵ^3 (Unit C) in terms of ϵ^1 (Unit A) and ϵ^2 (Unit B) so that the capital equality constraint is maintained.

Solution. Plan capital is the difference between assets and plan premium (Equation 1.1). Given the event 10 assumption, both scale linearly in ϵ , so Q does as well. The requirement is

$$47.8 = \sum_i Q_p^i (1 + \epsilon^i) = 2.1(1 + \epsilon^1) + 1.3(1 + \epsilon^2) + 44.4(1 + \epsilon^3),$$

therefore

$$\epsilon^3 = -\frac{2.1}{44.4}\epsilon^1 - \frac{1.3}{44.4}\epsilon^2 = -0.047297\epsilon^1 - 0.029279\epsilon^2. \quad \square$$

A naive approach to optimization might take the linear approximation at face value and take the biggest and most advantageous change in scale parameters possible. Let's see how this plays out with InsCo.

- The most advantageous first move is to reduce Unit A by 50%. This increases EVA by $(-0.2092) \cdot (-0.5) = 0.1046$. It also changes capital by $(2.1) \cdot (-0.5) = -1.05$.
- Now we need to increase overall writings to use up all available capital and get it back to the target. Unit B has a positive EVA/capital ratio and Unit C a negative, so Unit B has the most advantage. We would have to increase Unit B's scale by $1.05/1.3 = 0.808$ to completely use the capital, but this is outside the constrained range. Scaling by the max allowed 50% adds $1.3 \cdot 0.5 = 0.65$ to the capital, bringing the net position to $-1.05 + 0.65 = -0.4$. It also brings the EVA up by $0.0625 \cdot 0.5 = 0.0313$ to 0.1358.
- We need a further adjustment, and the only candidate left is Unit C. Here we need a scale change of $0.4/44.4 = 0.0090$, which is well within the allowed range. (The formula from the previous exercise tells us the same thing.) That brings the change in capital back to zero but changes EVA by $(-1.2186) \cdot 0.0090 = -0.0110$ for a net EVA of -1.2403 and an improvement of 0.1249. Adding these predicted changes to the original figures gives the results in Table 5.6.

Table 5.6. InsCo linear predictions for EVA and capital.

Quantity	Unit A	Unit B	Unit C	Portfolio
Scale change ϵ^i	-0.5	0.5	0.009	
Plan premium P_p	6.9500	28.0500	19.7766	54.7766
Required premium P_R	7.0546	27.9562	21.0061	56.0169
EVA $V = P_p - P_R$	-0.1046	0.0938	-1.2295	-1.2403
Required assets a	8.0000	30.0000	64.5766	102.5766
Plan capital $Q_P = a - P_p$	1.0500	1.9500	44.8000	47.8000

Now, had these moves been limited to, say 2%, we might be satisfied to stop here. In a real application, such a schedule of scale changes would be scrutinized for other business factors and questions of implementation feasibility. Some targets would probably be modified; for example, the Unit C scale change would no doubt be rounded to 1%.

However, with 50% scale changes, we need to be wary of nonlinear effects. Let us rerun the model at the new scales.

Our starting point is the original, 10-event loss distribution from Table 2.1. We multiply the losses in each column i by the scale factor $1 + \epsilon^i$, retabulate the portfolio total loss, and then sort on the portfolio loss. We apply the same asset metric—essentially the

maximum loss—and the same Wang transform for pricing. We do *not* compute a new Wang parameter to obtain a 15% portfolio return. Rather, we use the same parameter because this represents the investors' risk appetite. With a new loss distribution and new asset requirement, we do not expect the same required return. The new loss table is set out in Table 5.7. The first column shows the original sorted event number j , highlighting the reordering that has occurred with the change in volume by unit; see Table 2.1. Events 3, 4, 5, 7, and 8 all move and are therefore given a different weight by the Wang distortion. This extensive reshuffling accounts for the large (relative to EVA) change in required premium shown in Table 5.8.

Table 5.7. InsCo losses after rescaling and re-sorting.

j	p	Unit A	Unit B	Unit C	Portfolio	S	$g(S)$	q
	0	0	0	0	0	1.0	1.000	0.000
1	0.1	7.5	10.5	0	18.0000	0.9	0.931	0.069
2	0.1	7.5	19.5	0	27.0000	0.8	0.851	0.080
7	0.1	7.5	24.0	9.0811	40.5811	0.7	0.766	0.086
5	0.1	6.5	30.0	7.0631	43.5631	0.6	0.675	0.091
3	0.1	2.5	30.0	11.0991	43.5991	0.5	0.579	0.096
6	0.1	2.5	40.5	8.0721	51.0721	0.4	0.479	0.101
8	0.1	13.0	28.5	10.0901	51.5901	0.3	0.373	0.106
4	0.1	3.5	49.5	0	53.0000	0.2	0.261	0.112
9	0.1	8.5	12.0	40.3604	60.8604	0.1	0.140	0.121
10	0.1	8.0	30.0	64.5766	102.5766	0	0.000	0.140

Table 5.8. InsCo model rerun, EVA, and capital solutions.

Quantity	Unit A	Unit B	Unit C	Portfolio
Scale change ϵ^i	-0.5000	0.5000	0.0090	
Plan premium P_p	6.9500	28.0500	19.7766	54.7766
Required premium P_R	6.8115	28.6040	20.8271	56.2426
EVA $V = P_p - P_R$	0.1385	-0.5540	-1.0505	-1.4660
Required assets a	8.0000	30.0000	64.5766	102.5766
Plan capital $Q_p = a - P_p$	1.0500	1.9500	44.8000	47.8000
EVA/capital ratio $r = V/Q_p$	0.1319	-0.2841	-0.0234	

The final EVA of -1.466 is not only not as good as the predicted -1.240; it is *worse* than the original -1.365. Such is the curse of nonlinearity. Studying the derivatives here, it seems we need to go backwards, increasing the scale of Unit A and decreasing Unit B.

If we run the model with Unit B rescaled at many points throughout the $[-0.5, 0.5]$ range, we obtain the results shown in Figure 5.4. Nonlinearity is obvious. This is why most optimization algorithms take relatively small steps when updating parameters.

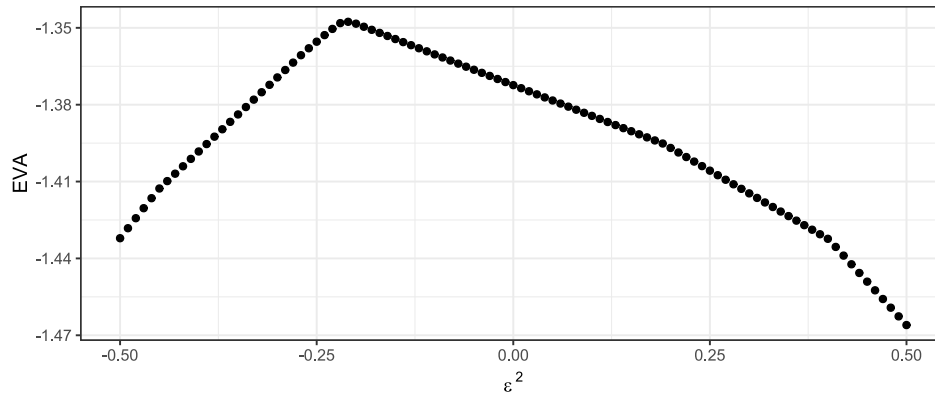


Figure 5.4. EVA at -50% Unit A and various Unit B.

Applying the Nelder-Mead Simplex algorithm, we obtain the surprising result that the scale constraints are not binding at the optimal solution. Unit A is to be reduced by almost, but not quite, 50%. Unit B is to be reduced by almost 21%. Details are in Table 5.9.

Table 5.9. InsCo model rerun, EVA, and capital solutions.

Quantity	Unit A	Unit B	Unit C	Portfolio
Scale ϵ^i	-0.478165	-0.209790	0.028758	
Plan premium P_p	7.253507	14.776927	20.163665	42.194098
Required premium P_R	7.269164	14.816395	21.453658	43.539217
EVA $V = P_p - P_R$	-0.015658	-0.039468	-1.289993	-1.345119
Required assets a	8.349360	15.804200	65.840538	89.994098
Plan capital $Q_p = a - P_p$	1.095854	1.027273	45.676874	47.800000
EVA/capital ratio $r = V/Q_p$	-0.014288	-0.038420	-0.028242	-0.080950

A grid search shows the situation in the neighborhood of the solution. See Figure 5.5. The optimal point has an EVA value of -1.345 . The innermost contour represents an EVA level set of -1.350 . Parameters inside this contour get at least 75% of the way from the original EVA to the optimal EVA.

The solution suggests that Unit A and Unit B both need further reduction because of negative EVA. However, there is more going on here. Not only is there nonlinearity, there is discontinuity in first derivatives as evidenced by Figure 5.4. If we were to change the Unit B scale ever so slightly, from -0.20979 to -0.209795 , we obtain dramatically different EVA allocations, shown in Table 5.10.

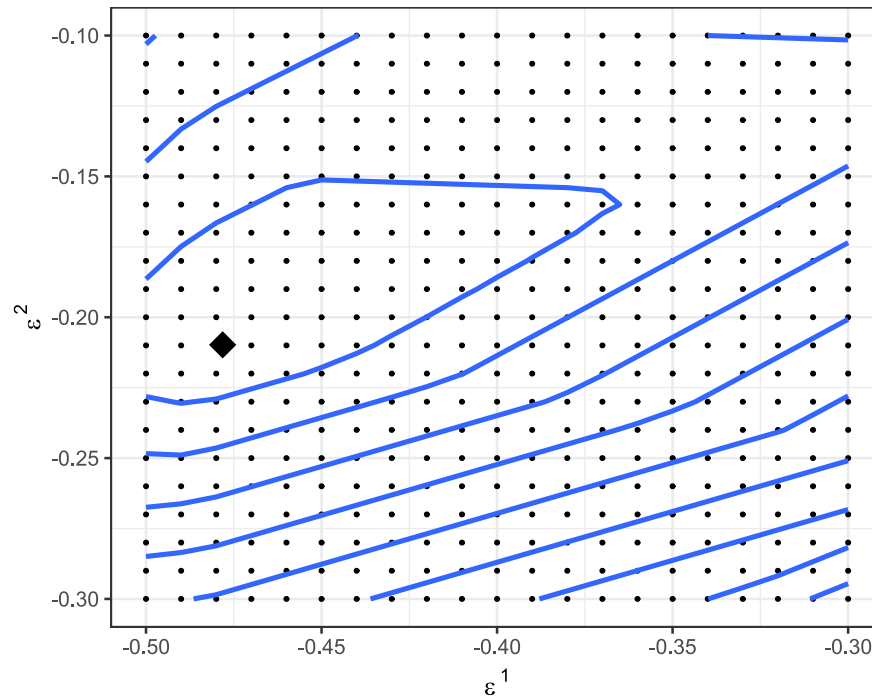


Figure 5.5. Contours of EVA at various Unit A and Unit B.

Table 5.10. InsCo model rerun, EVA, and capital solutions.

Quantity	Unit A	Unit B	Unit C	Portfolio
Scale ϵ^i	-0.478165	-0.209795	0.028759	-0.659201
Plan premium P_p	7.253507	14.776834	20.163668	42.194008
Required premium P_R	7.334042	14.607535	21.597550	43.539127
EVA $V = P_p - P_R$	-0.080535	0.169298	-1.433882	-1.345119
Required assets a	8.349360	15.804100	65.840548	89.994008
Plan capital $Q_p = a - P_p$	1.095854	1.027267	45.676880	47.800000
EVA/capital ratio $r = V/Q_p$	-0.073491	0.164805	-0.031392	0.059922

5.5.3. Summary

The term “portfolio optimization” is quite ambitious. There are some questionable assumptions that have to be imposed on the problem to get to a solution. There are inherent nonlinearities and discontinuous derivatives. The uncertainties (see Section 8.1) in the inputs and the underlying model itself no doubt add breadth to the range of near-optimal solutions. One should approach the task with a healthy dose of humility.

Perhaps the best that can be hoped for is to provide directions for portfolio improvement that are probably, approximately, correct.

6. Selecting and calibrating an SRM

Assume we have a can opener.

—Unknown

In Chapter 1, we saw that the overall portfolio profitability target—whether stated as total premium, margin, or return on capital—is typically given as an input. In Chapter 4, we saw the crucial importance of the distortion function in determining how that profit requirement is distributed to portfolio units. This chapter discusses how to develop a suitable distortion function.

6.1. The envelope of possibilities and the five representatives

TVaR is an SRM. The distortion function corresponding to TVaR_p goes linearly from $g(0) = 0$ to $g(1 - p) = 1$ and then stays at 1 for $s > 1 - p$; see Figure 4.3. A weighted combination of SRMs with positive weights summing to one (i.e., a convex combination) is also an SRM. In particular, we call such a weighted sum of two TVaRs a **bi-TVaR**.

Exercise. Derive an explicit expression for the distortion function associated with the bi-TVaR $\theta \text{TVaR}_{p_0} + (1 - \theta) \text{TVaR}_{p_1}$.

For any portfolio (i.e., loss random variable) X , there are two extreme values at which SRMs can price it. At one extreme, $\text{TVaR}_0(X) = E[X]$ is the minimum possible price; at the other, $\text{TVaR}_1(X) = \max(X)$ is the maximum.

Exercise. Prove that for any price P between $E[X]$ and $\max(X)$, there are infinitely many SRMs ρ with $\rho(X) = P$.

Solution. Because $\text{TVaR}_p(X)$ is continuous in p , the intermediate value theorem implies that there is a p^* with $\text{TVaR}_{p^*}(X) = P$. Similarly, because $\rho_\theta(X) \equiv \theta \text{TVaR}_0 + (1 - \theta) \text{TVaR}_1$ is continuous in θ , there is a θ^* with $\rho_{\theta^*}(X) = P$. The first is a TVaR and the second is a bi-TVaR, so they are distinct. Any convex combination of the two—and there are infinitely many of them—will also price X as P . \square

There are more than that, however. The same logic applies to any bi-TVaR $\theta \text{TVaR}_{p_0} + (1 - \theta) \text{TVaR}_{p_1}$ with $p_0 < p^* < p_1$; there is a θ^* pricing it at P . There are infinitely many of those. Combinations of those also work. Combinations of those combinations also work. If we pass to the limit and construct infinite combinations, we get SRMs with smooth curves for distortion functions like the Wang transform.

So there are many, many SRMs ρ that have $\rho(X) = P$. But applied to a particular unit (via NA), they are unlikely to all give the same price to the unit. It would be useful to be able to explore the limits of high and low unit prices consistent with the given portfolio price. Mildenhall and Major (2022) give an in-depth theoretical explanation, with Section 11.2 detailing the construction. Fortunately, the five representative distortions introduced in Section 4.6 do a good job of spanning the space of possibilities from body- (volatility-)centric to tail- (extreme risk-)centric and so in this monograph, we will consider only them.

A warning, however: unless one distortion g dominates another h in the sense that $g(s) \geq h(s)$ for all s with strict inequality for some s , then there are risks X and Y such that $\rho_g(X) > \rho_g(Y)$ and $\rho_h(X) < \rho_h(Y)$, i.e., g and h disagree about the relative price (riskiness) of X and Y . That is, it is not generally possible to put distortion functions in an unambiguous order by implied price across all risks.

Exercise. Verify that CCoC in Table 4.7 is the bi-TVaR between TVaR_1 and TVaR_0 with $\theta = 0.87$.

Solution. Table 6.1 shows the q values corresponding to $\text{TVaR}_0 = E[X]$ and $\text{TVaR}_1 = \max(X)$. The first consists of $q = p$, and the second consists of $q = 1$ for the top event only, zero elsewhere.

Table 6.1. q values for $\text{TVaR}_0 = E[X]$, $\text{TVaR}_1 = \max(X)$, and the $\theta = 0.87$ blend.

k	p	S	TVaR_0	TVaR_1	Blend
0	0	1	0	0	0
1	0.1	0.9	0.1	0	0.087
2	0.1	0.8	0.1	0	0.087
3	0.1	0.7	0.1	0	0.087
4	0.4	0.3	0.4	0	0.348
5	0.1	0.2	0.1	0	0.087
6	0.1	0.1	0.1	0	0.087
7	0.1	0	0.1	1	0.217

The final column, a 1:0.15 blend of the previous two, matches the CCoC column in Table 4.7. □

6.2. Required premium by distortion

Table 6.2 shows required premiums and loss ratios for each unit, by distortion function. Remember, *required* here means required to meet the internally specified 15% return on capital hurdle. Also shown are plan premiums and loss ratios.

Table 6.2. Example units with pricing by distortion.

View	Unit A		Unit B		Unit C		Portfolio	
	Premium	Loss Ratio	Premium	Loss Ratio	Premium	Loss Ratio	Premium	Loss Ratio
CCoC	13.739	97.5%	18.522	98.8%	21.304	69.9%	53.565	87.0%
Dual	14.127	94.9%	19.117	95.7%	20.321	73.3%	53.565	87.0%
PH	14.060	95.3%	18.349	99.7%	21.156	70.4%	53.565	87.0%
TVaR	13.783	97.2%	20.412	89.7%	19.371	76.9%	53.565	87.0%
Wang	14.109	95.0%	18.637	98.2%	20.818	71.6%	53.565	87.0%
Plan	13.9	96.4%	18.7	97.9%	19.6	76.0%	52.2	89.3%

Relative to model benchmarks, the entire portfolio is underpriced by $53.565 - 52.2 = 1.365$. But which units need to do better, and by how much?

Table 6.3 shows the plan loss ratio, the range of target loss ratios generated by the five distortions, and which distortion suggests the lowest premium for each unit. A higher required loss ratio implies a lower margin and premium. Unit management may regard a lower premium as better because it makes it easier to sell policies in the marketplace. The table shows that for each unit there is at least one distortion whose pricing it can meet, because the plan loss ratio is no greater than the distortion indicated loss ratio. For example, Unit B's management would be happy to target the PH-indicated 0.997 loss ratio instead of the current plan's 0.979. However, there is no single distortion whose pricing all units can meet. How should the modeler proceed? A monograph with two authors can offer two paths.

Table 6.3. Range of implied loss ratios and cheapest distortion by unit.

Unit	CCoC	PH	Wang	Dual	TVaR	Plan	Cheapest
Unit A	97.5%	95.3%	95.0%	94.9%	97.2%	96.4%	CCoC
Unit B	98.8%	99.7%	98.2%	95.7%	89.7%	97.9%	PH
Unit C	69.9%	70.4%	71.6%	73.3%	76.9%	76.0%	TVaR

Before discussing the two paths, it is worthwhile reflecting on how a market distortion emerges. Insurance is fundamentally risk sharing. The market consists of individuals with different endowments, owning different risks, and understood to have their own risk appetites. Each individual can cede risk into the pool and assume risk from it

as a shareholder or mutual company owner. Typically there are more of the former than the latter. We assume appetites for insurable, diversifiable risk are given by distortion risk measures. The appetite for investment risk, with its systematic components, may be different. The efficient insurance market solution is for everyone to pool their risks together, for maximal diversification benefit, and then to layer out the total risk to the individual offering the cheapest price for it in each layer. This corresponds to looking at the minimum of the individual distortion functions. Figure 4.3 could correspond to a market of five individuals. In it, the TVaR individual assumes large risks (above about the 75th percentile, lower left) and CCoC the rest. TVaR is risk neutral in the tail, and equity financing is well known to be an effective solution for nicely diversifiable risk. This theoretical model is described in a famous paper Jouini et al. (2008). There are many steps between a simple model and reality, but it is a good picture to bear in mind for the rest of this chapter. Readers who work for brokers may be happy to learn the model implies an important role for them matching risk to its cheapest financing. Finally, notice the model implies the risk appetite that matters is that of *investors*, not management—despite management often talking about “their risk appetite.”

6.3. Zeroing in on a distortion function

Path 1 holds there is a *best* distortion function that should be discoverable by the modeler. On this path, the modeler must avoid excessive naval gazing. Present the facts, being the range of indications from Table 6.2 with relevant commentary, but recognize the decision is for unit and corporate management to negotiate. As always, Table 6.2 hides many details and assumptions. What is the historical performance of each unit compared to prior plans? The growth prospects? The quality of the data and certainty in the simulations? Again, present facts and avoid selling against your model.

In an ideal world, we choose our distortion based on first principles and recommend each unit meets the target specified by that principle. A core principle of modern finance and accounting is to calibrate models to market observables where possible. Is the distortion consistent with cat bond pricing for extreme events? What about InsCo’s capital structure? A discussion with the finance department might be in order, to ask: What is the structure of the capital supporting the business and what are the costs of the various components of capital? For example, InsCo may have a Baa-rated corporate bond with initial annual default probability between 1.5% and 2% and a valuation equating to a risk-neutral default probability between 2.5% and 4%. These correspond to the exceedance probability s and distorted probability $g(s)$, respectively, of a distortion function. Which calibrated distortions are consistent with these findings? A quick analysis reveals that the Dual and Wang distortions are the most likely candidates, with TVaR coming in a distant third place.

6.4. Embracing the range

Path 2 holds that searching for a single best distortion is an unanswerable metaphysical quest, but it suggests there is much to learn from the range of indications. While no choice is correct, some are logically inconsistent and should be avoided.

The choice of a distortion corresponds to a statement of risk appetite. The ordering CCoC, PH, Wang, Dual, and TVaR we use to show the five representative distortions ranks them from most concerned about tail (extreme VaR) losses to most concerned about attritional (near plan) losses. The CCoC, driven by the maximum loss, is obviously the most tail-centric. Conversely, TVaR is the most attritional loss averse. This ranking corresponds to the heights of the distortion functions plotted in Figure 4.3; remember that left on the horizontal axis is extreme losses. How does management talk about risk, for example, in quarterly analyst calls for stock companies? Often, their statements reveal they understand their investors are more concerned with attritional volatility. After a large catastrophe event, the market expects all players to have large losses, whereas a miss on attritional volatility suggests management is not pricing correctly and can cause disquiet among investors. Managements prodded by their investors to be more concerned with volatility typically find targets driven by TVaR or the Dual distortion to better align with their intuitions.

The range of implications across different distortions turns out to overlap with the range of the reinsurance pricing cycle, so these are not just academic considerations. For example, in Table 9.10, the model recommended maximum buy-price for reinsurance is 7.6 (a 46% loss ratio or higher) under CCoC but only 4.8 (73% loss ratio) under TVaR. Over a market cycle the loss ratio for catastrophe reinsurance can vary substantially, and whereas a CCoC view may recommend a buy in all markets, TVaR could be more circumspect. The authors have both worked with managements who have expressed exactly these concerns and who have disliked a CCoC-driven, industry-standard approach analysis as a consequence.

Knowing the range of indications means you can determine whether a plan loss ratio is consistent with some risk appetite (it is within the range). For our simple model, Table 6.3 shows they all are, but that is not always the case. A unit with plan outside the range can be criticized because it is inconsistent with *any* risk appetite; it should be investigated accordingly. A plan far outside the range needs especial scrutiny.

7. Evaluating models

*The trouble with most of us is that we would rather
be ruined by praise than saved by criticism.*

—Norman Vincent Peale

This chapter encompasses three distinct tasks:

1. Evaluating your particular capital model.
2. Evaluating a third-party model, e.g., regulatory or rating agency model.
3. Evaluating a modeling platform you might use to build your own model.

7.1. Evaluating your model

There are several phases to evaluating your own model.

First, does it meet its design specifications in terms of features and functionality? You will refer back to the requirements document you prepared per Section 1.3. (You *did* prepare a formal document, didn't you?) This serves as a checklist as you review the features of your model.

Second, like any software development project, you need to run a battery of tests to verify it is behaving—especially calculating—correctly. If you built the model from the ground up, you might have created tests concurrently with module development. Use those, and add some more.

Third is validation: does the model produce realistic results? Whereas the first two phases apply to all the modules listed in Section 1.5, the third applies only to the Business Operations module. Validation might consist of expert judgment—but beware echo-chamber effects. Ideally, however, you have access to historical data that can be used to run the model as if it had existed in times past. Then, its outputs can be compared with what really did come to pass. Components of validation include:

1. **Best estimates:** With inputs corresponding to the plan, do central values (means or medians) of key variables (cash flows and derived accounting values) calibrate to the plan? Do the output distributions seem reasonable?
2. **Variability:** Do basic statistics (coefficient of variation, skewness, quantiles, etc.) of outputs seem reasonable? Are you getting what you expect? Are they consistent with historical (your own or industry) experience?

3. **Back testing:** If historical replays are possible, compute where the historical values lie in terms of percentiles of the output distributions, i.e., p values. The p values should reasonably be draws from uniform distributions. They should not cluster together at the low or high end—this indicates bias. They should not clump together in the middle—this indicates that historical variation is low compared to the model. Nor should they create a void in the middle—this indicates too much variation. Be careful not to overthink this, however. With only a handful of observations, only egregiously bad results can be considered statistically significant.
4. **Repeat:** As the model gets used, repeat back testing regularly (quarterly or annually). Build a growing library of p values to demonstrate that model ranges are wide enough and symmetrical.

7.2. Evaluating third-party models

Third-party models, if they include Business Operations simulation, should be evaluated—to the extent possible—in the manner described in Section 7.1.

Some third-party models, like minimum capital requirements created and used by regulators and rating agencies, do not simulate business outcomes but rather proceed directly from business characteristics as input to required capital as output. You may want to use such a model in one of several ways:

1. It is your only model; all you want to know is required capital. Unfortunately, this may not give you a clear path to allocating cost of capital and required premiums. However, it is a start if you have no capital model at all. Apply the non-Business Operations evaluation steps outlined in Section 7.1.
2. You intend to run it in parallel with your own model and compare results. Evaluate as in the previous bullet.
3. You intend to integrate it with your own model. This can mean several things.
 - a. Integrated reporting but otherwise separate calculations can be treated as previously but with additional testing to ensure the reporting is working correctly.
 - b. If the third-party model is to function as the Capital Adequacy module, then careful testing and review is needed to determine whether its outputs will be consistent with the workings of the Business Operations and Pricing and Allocation modules. For example, if a change to the portfolio moves the 95% and 99% loss quantiles upwards but the required capital goes down, then you might rethink the wisdom of relying solely on the third-party model. If it has capital allocation functionality, is it implementing something like CCoC? Can it be modified to perform an NA, or can that part be bypassed in favor of a different Pricing and Allocation module?

Some companies manage to (or are at least acutely aware of) their binding rating agency's or regulator's view of capital. You may be asked the point of an internal model,

given this binding view. The answer should go back to the charter for the internal model—after all, it is not a surprise that an external model is binding. What additional value was anticipated from an internal model *before you started building*? Reconsider why you are building an internal model if you cannot answer this question ahead of time.

7.3. Evaluating modeling platforms

Model building *platforms* range from programming languages with integrated development environments to heavily graphical user interface-based wiring diagram applications. What they have in common is that they are not ready-to-run models. How they differ is the extent to which they offer prefabricated computational modules and the support they offer for translating specifications into a model. Some programming languages, like R or Python, have many open-source libraries available to the public for free.

The first question to be addressed is: can the platform be used to build the model you have specified in Section 1.3? The second, trickier question is: how difficult will that be? Of course, budget limitations need also be taken into account. However, you *may* find a trade-off between direct expense to acquire the platform and labor expense to build the model with it.

Here are some items to consider when evaluating a modeling platform:

1. Modeling

- a. What accounting bases are supported for financial computations?
- b. How can loss payout patterns be modeled?
- c. How can assets be modeled (e.g., sweep accounts with proportionate investment strategies)?
- d. How can dependency be modeled (e.g., Iman-Conover shuffling, copulas, other)?
- e. What options are available for defining required capital?
- f. What options are available for defining and allocating required margins?

2. Parameters

- a. What parameters are explicitly variable and which are implicit and “hard-coded”?
- b. How are parameters stored? Can they easily be changed from one run to another?
- c. Is there an explicit provision for uncertainty in parameterization? (See also Section 8.1.)

3. Input and output

- a. What formats/databases are supported for data input? Report output? Simulation details?
- b. What are the report generation options?
- c. What tabular output options are there?
- d. What graphical output options are there?

4. Integration

- a. How is integration with Excel handled?
- b. Is integration with a separate Economic Scenario Generator possible?
- c. How is integration with commercial catastrophe models handled?

5. Running

- a. Are automatic multi-assumption model runs supported?
- b. Is there scripting? In what language?

6. System issues

- a. How are version control and model security handled?
- b. What are the hardware requirements (one PC, cluster, cloud)?
- c. Is there a front-end app or web interface?
- d. Is there an Application Programming Interface (API)?

7. Track record

- a. Are there models in production used for the London Market, Solvency II, and Swiss Solvency Test?
- b. What is the typical size of a client model (lines, year, assets, etc.) and what is the corresponding simulation run time?

8. Advanced topics

In this world, nothing is certain except death and taxes.

—Benjamin Franklin, letter to Jean-Baptiste Leroy

8.1. Dealing with uncertainty

The purpose of simulation is to estimate certain **quantities of interest** or **QoI**. A QoI is typically a quantile or other function (e.g., mean, risk measure) of the sample outcomes that the simulation created. Since it is being estimated from a sample, it cannot be measured with perfect accuracy because of sampling error. The more simulated samples are created, the more accurate the estimate can be. However, there is another reason that the estimate will not be accurate.

Parameter uncertainty here is to be distinguished from **process risk**. For example, when you simulate the outcome from a random variable with a lognormal distribution with particular parameters, you are modeling process risk. When you aren't sure what the parameter values should be, you are suffering from parameter uncertainty. The uncertainty around input parameters will propagate to uncertainty about the QoI. For example, it is well known that ignoring parameter uncertainty often systematically biases estimated quantiles downward. See Jewson et al. (2025) for an approach to adjust for this bias.

Parameter uncertainty arises from *nonsampling error*. There are many potential sources of this in a capital model. Data used to build the model could have erroneous entries, or it might not be of an appropriate vintage or refer to a population that is sufficiently relevant to the modeling task. In building the model, perhaps not all relevant factors have been taken into account. This is not an exhaustive list!

In addition, some parameters vary over time. A data snapshot taken to estimate parameters may be inappropriate to modeling the same phenomenon at a later date.

Parameter uncertainty will likely permeate your modeling efforts, especially if you are trying to make a model that can replicate historical experience. There are several ways to deal with uncertainty:

1. **Ignore it.** Use a best estimate of parameter values and let it go at that. While this may be acceptable for early versions of the model, it is not recommended for the long term. At the very least you must disclose that there is some uncertainty as to the QoI

values and what the business risks of that are. “The answer is A , but there is some uncertainty.”

2. **Report it.** Statistical analysis of data will usually give some indication of the range (e.g., variance) of parameter values consistent with a chosen model for the data. If you want to push your luck, that insight can be turned into a distributional model for the parameters. Rerun your model with carefully chosen exemplars or a sample from the putative distribution of those random parameters. Report the range or the resulting distribution of the QoI. “The answer is $A \pm \epsilon$.”
3. **Absorb it.** Classical (frequentist) statistics has prediction intervals to accommodate uncertainty in the parameters; Bayesian theory provides a formal framework for treating parameter uncertainty as another species of process risk. Essentially, this means treating the random parameter as mixing the distribution of QoI, creating what Bayesians call a predictive distribution; see Bolstad and Curran (2016) or Geisser (2017). This is explained in an actuarial context in Kreps (1997); however, Major (1999) finds reasons to doubt it is always the right approach. Prediction intervals are bigger than classically estimated intervals. The QoI is likely to be sensitive to the upper tail, which moves up. “The answer is $A + \delta$.”
4. **Isolate it.** Find a credible worst case to offer as an example of what can go wrong. This is the approach covered in Hansen and Sargent (2008) and Major et al. (2015). Again, this is likely to increase a risk-sensitive QoI. “The answer is A , but it could be as bad as $A + \gamma$.”

For a general treatment, see Smith (2024).

8.2. The Iman-Conover method

Correlation is not causation, but you can cause some correlation.

The Iman-Conover method is an easy way to induce correlation between marginal distributions. Here is the idea. Given samples of n values from r known marginal distributions X_1, \dots, X_r and a desired $r \times r$ (linear) correlation matrix Σ between them, reorder the i th sample to have the same rank order as the i th column of a reference distribution of size $n \times r$ with linear correlation Σ . The reordered output will then have the same rank correlation as the reference, and since linear correlation and rank correlation are typically close, it will have close to the desired correlation structure. What makes the Iman-Conover method work so effectively is the existence of easy algorithms to determine samples from reference distributions with prescribed correlation structures.

The Iman-Conover method is very powerful, quick, and easy to use. However, it is not magic, and it is constrained by mathematical reality. For given pairs of distributions there is often an upper bound on the *linear* correlation between them, so if you ask for very high correlation you may be disappointed. You should go back and ask yourself why you think the linear correlation is so high.

Iman-Conover is usually used to induce a desired linear correlation. There are other measures, such as rank correlation and Kendall's tau, that could be used (Mildenhall 2005) and the user should reflect on which is most appropriate to their application. If your target is rank correlation, Iman-Conover will give you an exact match.

The `aggregate` documentation¹ provides a detailed description of the algorithm, including discussion of using different reference distributions (called scores) in place of the normal, and a comparison with the normal copula method.

8.2.1. Loss or loss ratio?

Are you modeling correlation of *loss ratios* or correlation of *losses*? It is a surprising empirical fact that the former is often greater than the latter, leading to the conclusion that correlation in *underwriting results* is driven by correlation in *premium* more than loss; see Mildenhall (2022). These findings highlight a stark property-casualty split. Property tends to be loss-volatility driven, with loss correlation caused (and explained) by geographic proximity. Casualty tends to be premium-volatility driven, with loss ratio correlation caused by (account) underwriting and the underwriting cycle.

8.2.2. Example 1: A simple example

This section presents a simple example applying Iman-Conover to realistic distributions, using the `aggregate` library. It shows how easy it is to use Iman-Conover in modern programming languages. Another example, with more the details, appears in the `aggregate` documentation.² See Section 9.3 for instructions to set up your Python environment to replicate the code samples. The full source code to replicate the examples is available on the monograph website.³

The example works with $r = 3$ different marginals: Auto, general liability (GL), and Property. Losses are simulated using frequency-severity aggregate distributions with a lognormal severity and per occurrence limits. Frequency is a gamma-mixed Poisson (negative binomial) with varying CVs. See Mildenhall (2023b) for a description of the `aggregate` DecL language, used to specify these portfolios.

```
port ICExample
agg Auto      5000 loss 2000 xs 0 sev lognorm 30 cv 2 mixed gamma 0.2
agg GL        2000 loss 5000 xs 0 sev lognorm 50 cv 5 mixed gamma 0.3
agg Property  7000 loss 10000 xs 0 sev lognorm 100 cv 5 mixed gamma 0.1
```

This example uses only 1,000 simulations, but the software can easily handle 100,000 or more. The slowest part of the Iman-Conover algorithm is creating the scores. These are cached, so subsequent runs are very fast, even if the desired correlation matrix changes.

¹ https://aggregate.readthedocs.io/en/latest/5_technical_guides/5_x_iman_conover.html#algorithm.

² https://aggregate.readthedocs.io/en/latest/5_technical_guides/5_x_iman_conover.html#simple-example-of-iman-conover.

³ <https://casact.github.io/capital-modeling/>.

Table 8.1 shows summary statistics and the first few samples of the underlying multivariate distribution, and Figure 8.1 shows histograms of each marginal and scatter plots. By construction, the units are approximately independent.

Table 8.1. Set up and summary statistics of sample for Iman-Conover example using 1,000 samples.

(a) Summary statistics				(b) First 10 samples			
Index	Auto	GL	Property	Index	Auto	GL	Property
Count	1,000	1,000	1,000	0	4,970	207	5,684
Mean	5,011.7	2,092.4	7,117.8	1	4,882	1,599	9,050
SD	1,289.3	1,381.2	3,323.0	2	4,192	2,576	5,538
Min	1,375	145	1,524	3	6,569	2,107	6,865
25%	4,102.5	1,152.5	4,875.5	4	6,236	2,472	5,342
50%	4,914.5	1,773.5	6,412.5	5	4,206	1,835	5,738
75%	5,771.5	2,670.5	8,597.5	6	5,537	1,313	6,217
Max	9,486	12,537	27,108	7	4,483	4,720	12,113
CV	0.257	0.660	0.467	8	4,325	1,323	21,139
				9	5,241	541	22,005

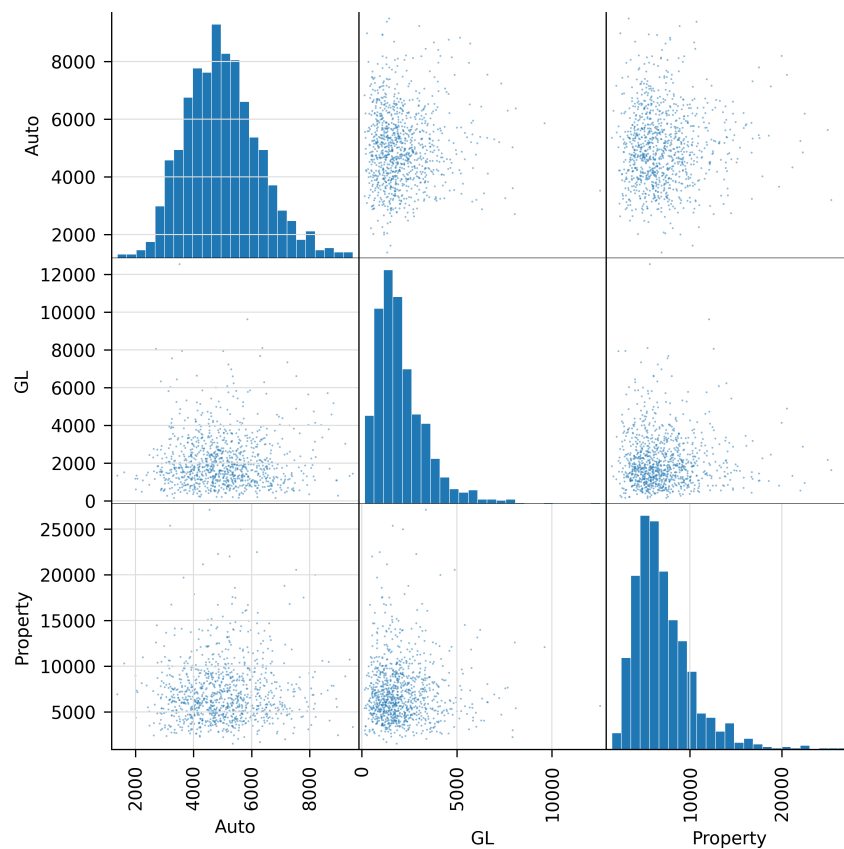


Figure 8.1. Uncorrelated marginals input to Iman-Conover method. Diagonal shows univariate histograms.

In addition to marginal distributions, Iman-Conover requires a target desired correlation matrix as input. This is produced by a separate analysis, and here we assume it is given by Table 8.2. Remember, a correlation matrix must be symmetric (easy to check) and positive definite (harder). Matrices computed from statistical analysis may fail to be positive definite if they are not computed from a multivariate sample. Iman-Conover fails if given a nonpositive definite matrix.

Table 8.2. Target correlation matrix.

Index	Auto	GL	Property
Auto	1	-0.3	0
GL	-0.3	1	0.8
Property	0	0.8	1

The Iman-Conover method is implemented in `aggregate` by the function `iman_conover`, so the transformation is simply

```
df2 = iman_conover(df, desired)
```

where `desired` is the desired correlation matrix. Table 8.3 shows the achieved linear and rank correlation, and the absolute and relative errors. Finally, Figure 8.2 provides a scatter plot of the shuffled inputs, making the correlation evident. It is possible to obtain a correlated sample from a `Portfolio` object by passing in a correlation matrix: the code below extracts a sample and automatically applies Iman-Conover.

```
df = port.sample(n, desired_correlation=desired, keep_total=False)
```

Table 8.3. Achieved correlation matrix and output errors with Iman-Conover method.

(a) Linear correlation				(b) Rank correlation			
Index	Auto	GL	Property	Index	Auto	GL	Property
Auto	1	-0.271	0.003	Auto	1	-0.287	0.003
GL	-0.271	1	0.772	GL	-0.287	1	0.779
Property	0.003	0.772	1	Property	0.003	0.779	1

(c) Absolute error				(d) Relative error			
Index	Auto	GL	Property	Index	Auto	GL	Property
Auto	0	0.029	0.003	Auto	0	-0.095	Infinite
GL	0.029	0	0.028	GL	-0.095	0	-0.035
Property	0.003	0.028	0	Property	Infinite	-0.035	0

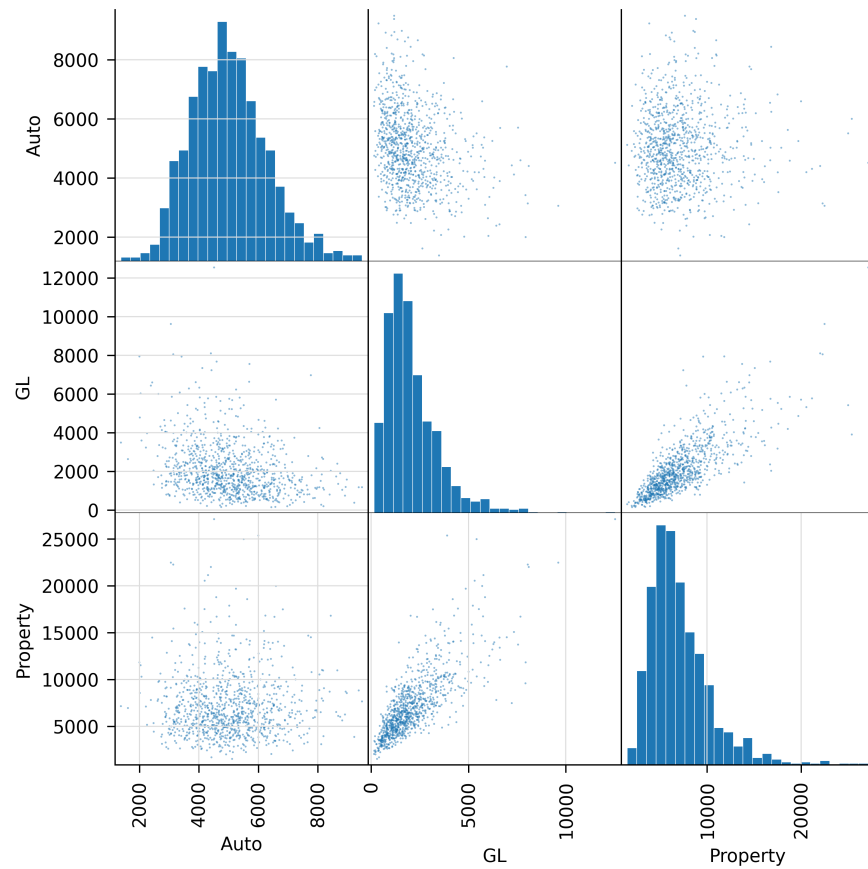


Figure 8.2. Correlated marginals output by the Iman-Conover method. Diagonal shows unchanged univariate histograms, but the scatter plots clearly show the introduction of correlation.

8.2.3. Different reference distributions

The Iman-Conover algorithm relies on a sample from a multivariate reference distribution called scores. A straightforward method to compute a reference is to use the Cholesky decomposition method applied to certain independent scores (samples with mean zero and standard deviation one). The standard algorithm uses standard normal scores. However, nothing prevents the use of other distributions for the scores, provided they are suitably normalized to have mean zero and standard deviation one. Several other families of multivariate distributions could be used, such as elliptically contoured distributions (which include the normal and t as a special cases) and the multivariate Laplace distribution. These are all easy to simulate.

The `aggregate` implementation provides t -distribution scores. It includes an argument `dof`. By default `dof = 0` resulting in a normal reference. A `dof > 0` is interpreted as degrees of freedom, and the reference becomes a t -distribution. Figure 8.3 reruns Section 8.2.2 using a t reference with two degrees of freedom, a very extreme choice. It results

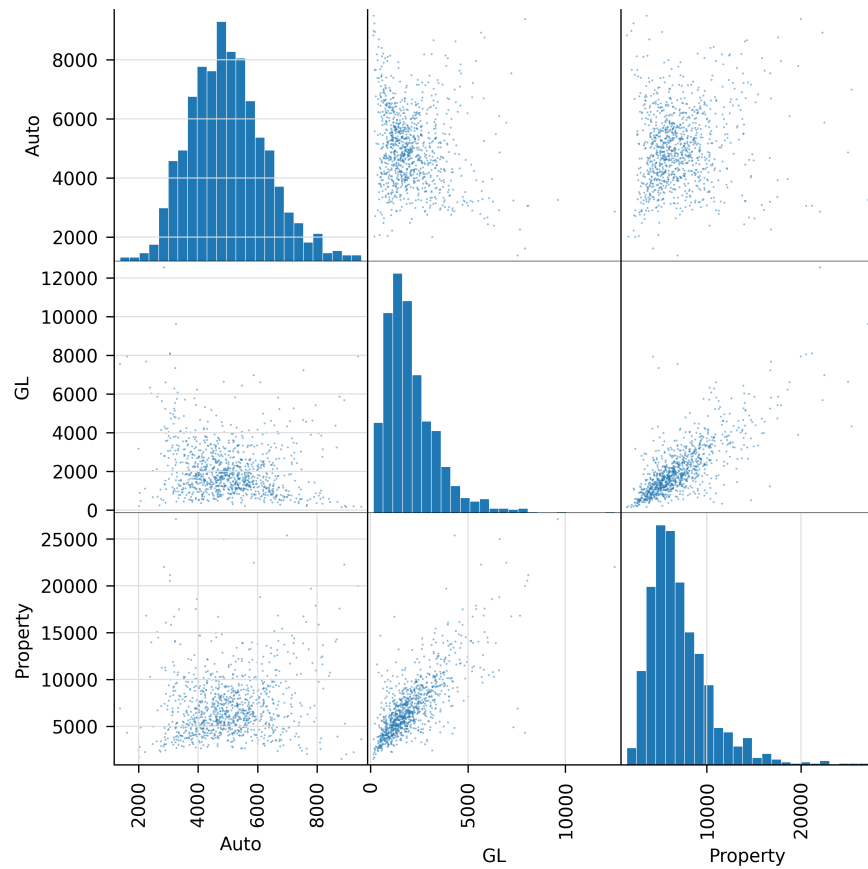


Figure 8.3. Correlated marginals using a t -distribution with two degrees of freedom.

in more tail correlation, evident in the pinched scatter plots. Be warned, however, that this modeling flexibility definitely falls into the “just because you can, doesn’t mean you should” category!

8.2.4. *Copulas as reference distributions*

The reader may be wondering about using a copula as a reference score. Alas, most copulas are only bivariate. Multivariate copulas, aside from those related to elliptical or Laplace distributions already mentioned, are Archimedean, which means they have a single parameter and cannot be calibrated to a desired correlation matrix.

See the `aggregate` documentation⁴ and Mildenhall (2005) for more details.

8.2.5. *Iman-Conover compared with the normal copula method*

The normal copula method is described in Wang (1998). Given a set of risks (X_1, \dots, X_k) with marginal cumulative distribution functions F_i and Kendall’s tau $\tau_{ij} = \tau(X_i, X_j)$ or rank correlation coefficients $r(X_i, X_j)$, it works as follows. Assume that (Z_1, \dots, Z_k) have

⁴ https://aggregate.readthedocs.io/en/latest/5_technical_guides/5_x_working_with_samples.html#alternative-scores.

a multivariate normal joint probability density function given by

$$f(z_1, \dots, z_k) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp(-z' \Sigma^{-1} z / 2),$$

$Z = (z_1, \dots, z_k)$, with correlation coefficients $\Sigma_{ij} = \rho_{ij} = \rho(Z_i, Z_j)$. Let $H(z_1, \dots, z_k)$ be their joint cumulative distribution function. Then

$$C(u_1, \dots, u_k) = H(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_k))$$

defines a multivariate uniform cumulative distribution function called the **normal copula**. Using H , the set of variables

$$X_1 = F_1^{-1}(\Phi(Z_1)), \dots, X_k = F_k^{-1}(\Phi(Z_k)) \quad (8.1)$$

have a joint cumulative function

$$F_{X_1, \dots, X_k}(x_1, \dots, x_k) = H(\Phi^{-1}(F_1(x_1)), \dots, \Phi^{-1}(F_k(x_k)))$$

with marginal cumulative distribution functions F_1, \dots, F_k . The multivariate variables (X_1, \dots, X_k) have Kendall's tau

$$\tau(X_i, X_j) = \tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin(\rho_{ij})$$

and Spearman's rank correlation coefficients

$$\text{rkCorr}(X_i, X_j) = \text{rkCorr}(Z_i, Z_j) = \frac{6}{\pi} \arcsin(\rho_{ij}/2),$$

so ρ can be adjusted to achieve the desired correlation.

In the normal copula method, we simulate from H and then invert using Equation 8.1. In the Iman-Conover method with normal scores we produce a sample from H such that $\Phi(z_i)$ are equally spaced between zero and one, and then, rather than invert the distribution functions, we make the j th order statistic from the input sample correspond to $\Phi(z) = j/(n+1)$ where the input has n observations. Because the j th order statistic of a sample of n observations from a distribution F approximates $F^{-1}(j/(n+1))$, we see the normal copula and Iman-Conover methods are doing essentially the same thing.

While the normal copula method and the Iman-Conover method are confusingly similar, there are some important differences to bear in mind. Comparing and contrasting the two methods should help clarify how the two algorithms are different.

1. Wang (1998) shows the normal copula method corresponds to the Iman-Conover method when the latter is computed using normal scores and the Choleski trick.

2. The Iman-Conover method works on a given sample of marginal distributions. The normal copula method generates the sample by inverting the distribution function of each marginal as part of the simulation process.
3. Through the use of scores, the Iman-Conover method relies on a sample of normal variables. The normal copula method could use a similar method, or it could sample randomly from the base normals. Conversely, a sample could be used in the Iman-Conover method.
4. Only the Iman-Conover method has an adjustment to ensure that the reference multivariate distribution has exactly the required correlation structure.
5. Iman-Conover method samples have rank correlation exactly equal to a sample from a reference distribution with the correct linear correlation. Normal copula samples have approximately correct linear and rank correlations.
6. An Iman-Conover method sample must be taken in its entirety to be used correctly. The number of output points is fixed by the number of input points, and the sample is computed in its entirety in one step. Some Iman-Conover tools produce output, which is in a particular order. Thus, if you sample the n th observation from multiple simulations, or take the first n samples, you will not get a random sample from the desired distribution. However, if you select random rows from multiple simulations (or, equivalently, if you randomly permute the rows' output prior to selecting the n th), then you will obtain the desired random sample. It is important to be aware of these issues before using canned software routines.
7. The normal copula method produces simulations one at a time, and at each iteration the resulting sample is a sample from the required multivariate distribution. That is, output from the algorithm can be partitioned and used in pieces.

In summary, remember that these differences can have material, practical consequences, and it is important not to misuse Iman-Conover method samples.

8.2.6. *Block Iman-Conover (BIC)*

The Iman-Conover method can lead you into a state of sin: the sin of thinking you can reliably estimate a 100×100 correlation matrix (or larger). You can't; don't try—it's all noise. Look at the distribution of eigenvalues to see; start by computing the singular value decomposition (SVD) and compare your distribution with that of a random matrix. However, you might be able to say something about the correlation of chunks of your book. Generally (going back to premium is riskier than loss), lines that are underwritten together have correlated loss ratios, and properties that are nearby have correlated losses. The Block Iman-Conover (BIC) method is a way to use this information without fabricating large correlation matrices.

Description of BIC. BIC is best explained through an example. You underwrite auto liability, GL, and property in four regions: North, East, South, and West. Based on reliable studies, you have a sense of the loss (ratio) correlation between each line overall (a 3×3 matrix, with three parameters), and, separately, between the total regional results (a 4×4 matrix, with six parameters). The BIC produces a multivariate sample across all 12 line and region units that incorporates what you know, but using just these nine parameters rather than requiring 66 parameters (for a 12×12 matrix). Here's how it works. Pick a number of simulations n , then:

1. Produce four $n \times 3$ matrices with results by line for each region, R_1, R_2, R_3, R_4 .
2. Apply Iman-Conover to each R_i separately using your 3×3 between-line correlation matrix, to produce \tilde{R}_i .
3. Sum across simulations in each \tilde{R}_i to get total regional losses and assemble these into a $n \times 4$ matrix T .
4. Apply Iman-Conover to T using your 4×4 between-region correlation matrix, and capture the ranking.
5. Reorder each \tilde{R}_i by row according to the ranking in Step 4 to obtain $\tilde{\tilde{R}}_i$.
6. Assemble $\tilde{\tilde{R}}_i$ together into a $n \times 12$ multivariate sample X .

The result of this algorithm has the correct by-line correlation within each region—because in Step 5 rows are moved, which preserves their within-group correlation—and the correct correlation of total by region, from Step 4.

The `aggregate` package implements BIC as `block_iman_conover`. It takes arguments:

- `unit_losses`: A list of samples, in our example the four blocks R_i of losses by line within a region.
- `intra_unit_corrs`: A list of correlation matrices within (intra) each region.
- `inter_unit_corr`: The between (inter) unit correlation matrix.

The regions can have different numbers of lines, and the intra-correlation matrices can vary across regions. In the example below, we just use four regions with three lines, and we assume the intra-region by-line correlations are all the same.

8.2.7. Example 2: Multistate, multiline correlation

Example 2 implements the case described in Section 8.2.6: four regions each with three lines, and reliably estimated correlation matrices between lines and between regions. The next block shows the specification of the regional loss distributions in Decl.

```
port North
agg Auto_North 4000 loss 2000 xs 0 sev lognorm 30 cv 2 mixed gamma .2
```

```

agg GL_North      2000 loss  5000 xs 0 sev lognorm  50 cv 4 mixed gamma .3
agg Prop_North    2000 loss 10000 xs 0 sev lognorm 100 cv 5 mixed gamma .1

port East
agg Auto_East     3000 loss  2000 xs 0 sev lognorm  30 cv 2 mixed gamma .2
agg GL_East       1000 loss  5000 xs 0 sev lognorm  50 cv 4 mixed gamma .3
agg Property_East 3000 loss 10000 xs 0 sev lognorm 100 cv 5 mixed gamma .1

port West
agg Auto_West     5000 loss  2000 xs 0 sev lognorm  30 cv 2 mixed gamma .2
agg GL_West       2000 loss  5000 xs 0 sev lognorm  50 cv 4 mixed gamma .3
agg Property_West 2000 loss 10000 xs 0 sev lognorm 100 cv 5 mixed gamma .1

port South
agg Auto_South    5000 loss  2000 xs 0 sev lognorm  30 cv 2 mixed gamma .2
agg GL_South      2000 loss  5000 xs 0 sev lognorm  50 cv 4 mixed gamma .3
agg Prop_South    7000 loss 10000 xs 0 sev lognorm 100 cv 5 mixed gamma .1

```

The losses by line are simulated using frequency-severity aggregate distributions with a lognormal severity and per occurrence limits. Frequency is a gamma-mixed Poisson (negative binomial) with varying CVs. See Mildenhall (2023b) for a description of the aggregate DecL language, used to specify these portfolios. The South region portfolio was previously used to illustrate the standard Iman-Conover method in Section 8.2.2.

Table 8.4 shows the expected losses, modeling error, CV, and 99th percentile of total loss by region. Model error means the difference between the output of the Fast Fourier Transform algorithm and the requested input mean losses. As you can see, the two are almost identical. The CVs vary reflecting volume and mix of business between regions.

Table 8.4. Expected losses, CV, and 99th percentile loss by region.

Region	Expected Loss	Model Error	CV	99th Percentile
North	8,000.0	−0.000	0.298	16,237
East	7,000	−0.000	0.347	15,920
South	14,000	−0.000	0.266	25,737
West	9,000	−0.000	0.276	17,331

Next, we sample $n = 1,000$ losses from each region with the intra-region correlation. In this case, all the intra-region by-line correlations are the same, but the method allows them to vary.

The BIC is then run using:

```

ans = block_iman_conover(region_losses, intra_region_correl,
    inter_region_correl, True)

```

The three panels of Table 8.5 (respectively, Table 8.6) show the target and achieved inter-region (respectively, intra-region) correlation and the error (difference). The achieved results are very close to those desired.

Table 8.5. Target and achieved inter-region correlations.

(a) Target inter-region correlation

Index	North	East	South	West
North	1	0.7	0.2	0
East	0.7	1	0.4	0.2
South	0.2	0.4	1	0.8
West	0	0.2	0.8	1

(b) Achieved inter-region correlation

Index	North	East	South	West
North	1	0.683	0.167	-0.002
East	0.683	1	0.369	0.197
South	0.167	0.369	1	0.799
West	-0.002	0.197	0.799	1

(c) Difference

Index	North	East	South	West
North	0	0.017	0.033	0.002
East	0.017	0	0.031	0.003
South	0.033	0.031	0	0.001
West	0.002	0.003	0.001	0

Table 8.6. Achieved (a-d) and target (e) intra-region correlation.

(a) North

Index	N-AL	N-GL	N-Pr
N-AL	1	0.579	0.180
N-GL	0.579	1	0.332
N-Pr	0.180	0.332	1

(b) East

Index	E-AL	E-GL	E-Pr
E-AL	1	0.544	0.191
E-GL	0.544	1	0.318
E-Pr	0.191	0.318	1

(c) South

Index	S-AL	S-GL	S-Pr
S-AL	1	0.571	0.202
S-GL	0.571	1	0.354
S-Pr	0.202	0.354	1

(d) West

Index	W-AL	W-GL	W-Pr
W-AL	1	0.580	0.176
W-GL	0.580	1	0.330
W-Pr	0.176	0.330	1

(e) Target

Index	W-AL	W-GL	W-Pr
W-AL	1	0.6	0.2
W-GL	0.6	1	0.4
W-Pr	0.2	0.4	1

Table 8.7 shows the overall correlation matrix at the line-region level (12×12 matrix), and Figure 8.4 shows an image plot of the same matrix. The underlying block structure is clear from the image.

Table 8.7. Achieved correlation across lines and regions.

Index	N-AL	N-GL	N-Pr	E-AL	E-GL	E-Pr	S-AL	S-GL	S-Pr	W-AL	W-GL	W-Pr
N-AL	1	0.579	0.180	0.319	0.285	0.418	0.072	0.042	0.115	-0.011	-0.071	0.008
N-GL	0.579	1	0.332	0.354	0.391	0.495	0.111	0.076	0.143	0.010	-0.025	0.010
N-Pr	0.180	0.332	1	0.281	0.363	0.426	0.129	0.089	0.067	-0.009	0.007	0.028
E-AL	0.319	0.354	0.281	1	0.544	0.191	0.170	0.097	0.216	0.088	0.122	0.099
E-GL	0.285	0.391	0.363	0.544	1	0.318	0.127	0.122	0.255	0.079	0.148	0.130
E-Pr	0.418	0.495	0.426	0.191	0.318	1	0.240	0.261	0.237	0.089	0.117	0.132
S-AL	0.072	0.111	0.129	0.170	0.127	0.240	1	0.571	0.202	0.300	0.342	0.363
S-GL	0.042	0.076	0.089	0.097	0.122	0.261	0.571	1	0.354	0.385	0.431	0.423
S-Pr	0.115	0.143	0.067	0.216	0.255	0.237	0.202	0.354	1	0.541	0.579	0.506
W-AL	-0.011	0.010	-0.009	0.088	0.079	0.089	0.300	0.385	0.541	1	0.580	0.176
W-GL	-0.071	-0.025	0.007	0.122	0.148	0.117	0.342	0.431	0.579	0.580	1	0.330
W-Pr	0.008	0.010	0.028	0.099	0.130	0.132	0.363	0.423	0.506	0.176	0.330	1

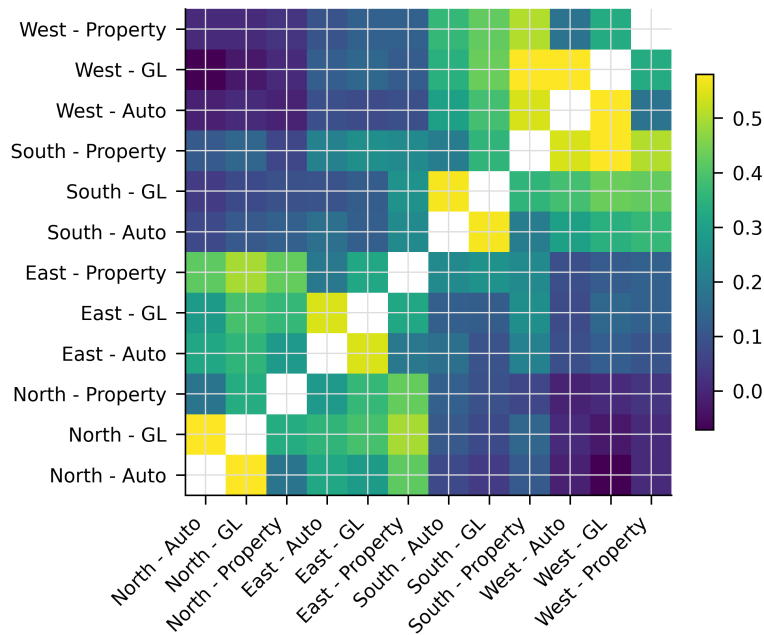


Figure 8.4. Image plot of achieved intra-region correlation. The block structure is clearly evident.

8.2.8. Theoretical derivation of the Iman-Conover method

Suppose that \mathbf{M} is an n -element sample from an r -dimensional multivariate distribution, so \mathbf{M} is an $n \times r$ matrix. Assume that the columns of \mathbf{M} are uncorrelated, have mean zero,

and have a standard deviation of one. Let M' denote the transpose of M . These assumptions imply that the correlation matrix of the sample M can be computed as $n^{-1}M'M$, and because the columns are independent, $n^{-1}M'M = \text{id}$. (There is no need to scale the covariance matrix by the row and column standard deviations because they are all one. In general $n^{-1}M'M$ is the covariance matrix of M .)

Let S be a correlation matrix, i.e., S is a positive semidefinite symmetric matrix with ones on the diagonal and all elements ≤ 1 in absolute value. In order to rule out linearly dependent variables, assume S is positive definite. These assumptions ensure S has a Choleski decomposition

$$S = C'C$$

for some upper triangular matrix C ; see Golub and Van Loan (2013) or Press et al. (1992). Set $T = MC$. The columns of T still have mean zero, because they are linear combinations of the columns of M , which have zero mean by assumption. It is less obvious, but still true, that the columns of T still have standard deviation one. To see why, remember that the covariance matrix of T is

$$n^{-1}T'T = n^{-1}C'M'MC = C'C = S,$$

since $n^{-1}M'M = \text{id}$ is the identity by assumption. Now S also is actually the correlation matrix because the diagonal is scaled to standard deviation one, so the covariance and correlation matrices coincide. The process of converting M , which is easy to simulate, into T , which has the desired correlation structure S , is the theoretical basis of the Iman-Conover method.

It is important to note that estimates of correlation matrices, depending on how they are constructed, need not have the mathematical properties of a correlation matrix. Therefore, when trying to use an estimate of a correlation matrix in an algorithm, such as the Iman-Conover, which actually requires a proper correlation matrix as input, it may be necessary to check the input matrix does have the correct mathematical properties.

Next, we discuss how to make $n \times r$ matrices M , with independent, mean zero columns. The basic idea is to take n numbers a_1, \dots, a_n with $\sum_i a_i = 0$ and $n^{-1} \sum_i a_i^2 = 1$, use them to form one $n \times 1$ column of M , and then to copy it r times. Finally, randomly permute the entries in each column to make them independent as columns of random variables. Iman and Conover (1979) call these a_i scores. They discuss several possible definitions for scores, including scaled versions of $a_i = i$ (ranks) and a_i uniformly distributed. They note that the shape of the output multivariate distribution depends on the scores. All of the examples in their paper use normal scores.

Given that the scores will be based on normal random variables, we can either simulate n random standard normal variables and then shift and rescale to ensure mean zero

and standard deviation one, or we can use a systematic sample from the standard normal, $a_i = \Phi^{-1}(i/(n+1))$. By construction, the systematic sample has mean zero, which is an advantage. Also, by symmetry, using the systematic sample halves the number of calls to Φ^{-1} . For these two reasons we prefer it in the algorithm below.

The correlation matrix of M , constructed by randomly permuting the scores in each column, will only be approximately equal to id because of random simulation error. In order to correct for the slight error that could be introduced, Iman and Conover use another adjustment in their algorithm. Let $EE = n^{-1}M'M$ be the actual correlation matrix of M , and let $EE = F'F$ be the Choleski decomposition of EE , and define $T = MF^{-1}C$. The columns of T have mean zero, and the covariance matrix of T is

$$\begin{aligned} n^{-1}T'T &= n^{-1}C'F'^{-1}M'MF^{-1}C \\ &= C'F'^{-1}EEF^{-1}C \\ &= C'F'^{-1}F'FF^{-1}C \\ &= C'C \\ &= S, \end{aligned}$$

and hence T has correlation matrix exactly equal to S , as desired. If EE is singular, then the column shuffle needs to be repeated.

Now that the reference distribution T with exact correlation structure S is in hand, all that remains to complete the Iman-Conover method is to reorder each column of the input distribution X to have the same rank order as the corresponding column of T .

8.3. Allocating capital, if you must

Within each layer, we can allocate its expected losses and premiums to the various portfolio units. Having done so, capital allocation to units follows.

Admonition. We can't emphasize strongly enough that you **should not allocate capital**. Different layers of capital do not have the same cost (think: equity, junk debt, government bonds), therefore you do not know the appropriate target return on your allocated capital. Without that, you can't say if a projected return is good or bad. Allocate margin, not capital!

Warnings about smoking don't work either. We know your CFO, senior management, and investors discuss allocating underwriting capacity in the language of allocating capital, and we previously admonished you to speak the language of your customers. This section describes two methods you can use to translate allocated margins into allocated capital.

The first method is obvious: simply assume all capital earns the weighted average cost of capital and allocate capital by dividing allocated margin by the weighted average cost of capital. This will result in an additive allocation of capital. This has the advantage of being simple to implement, but the disadvantage is that it obscures—nay, denies!—the fact that different segments of capital earn different returns.

The second method has a more theoretically sound basis. It applies the principles discussed in Chapter 4 to obtain an allocation of capital Q and unit-specific rates of return i^i consistent with the SRM approach.

Layer k is ΔX_k dollars wide but is triggered by any portfolio loss $X > X_k$. We propose that the expectation of that loss be allocated to units by equal priority, that is, in proportion to the expected losses of the units *in those triggering events*. We may define the layer- k **expected loss share** to unit i as

$$\alpha_k^i := \mathbb{E}_p \left[\frac{X^i}{X} \mid X > X_k \right].$$

In particular, this means

$$\alpha_k^i S_k = \mathbb{E}_p \left[\frac{X^i}{X} 1_{X > X_k} \right] = \sum_{j > k} \frac{X_j^i}{X_j} p_j. \quad (8.2)$$

The α^i sum to one across the units because (conditional) expectations are linear. The expected loss to layer k , $S_k \Delta X_k$, can thus be partitioned into $\alpha_k^i S_k \Delta X_k$ across the various units. The total portfolio loss, $\mathbb{E}[X]$, is allocated to unit i as

$$\mathbb{E}[X^i] = \sum_j \alpha_j^i S_j \Delta X_j. \quad (8.3)$$

Exercise. Prove Equation 8.3, given Equation 8.2.

Solution.

$$\begin{aligned} \sum_j \alpha_j^i S_j \Delta X_j &= \sum_{k=0}^{n-1} \left(\sum_{j=k+1}^n \frac{X_j^i}{X_j} p_j \right) (X_{k+1} - X_k) \\ &= \sum_{j=1}^n \frac{X_j^i}{X_j} p_j \sum_{k=0}^{j-1} (X_{k+1} - X_k) \\ &= \sum_{j=1}^n \frac{X_j^i}{X_j} p_j X_j \\ &= \sum_{j=1}^n X_j^i p_j \\ &= \mathbb{E}[X^i] \end{aligned}$$

noting $X_0 = X_0^i = 0$. □

What about premium? We propose that the premium associated with layer k be allocated to units in proportion to the *distorted* expected losses of the units in those triggering events. This leads to the definition of the **distorted expected loss share**

$$\beta_k^i := \mathbb{E}_q \left[\frac{X^i}{X} \mid X > X_k \right].$$

In particular, this means

$$\beta_k^i g(S_k) = \sum_{j>k} \frac{X_j^i}{X_j} q_j = \sum_{j>k} \frac{X_j^i Z_j}{X_j} p_j,$$

where, recall, $Z_j = q_j / p_j$ is the event weight.

The premium to layer k , given by $P_k = g(S_k) \Delta X_k$, is partitioned to the units as $\beta_k^i g(S_k) \Delta X_k$. The total premium is therefore allocated to the units as

$$\mathbb{E}[X^i Z(X)] = \sum_j \beta_j^i g(S_j) \Delta X_j$$

being the total premium to unit i .

Exercise. Prove the previous equality.

Solution. The same steps as the preceding exercise. □

Since margin is the difference between premium and expected loss (Equation 4.1), we immediately have the allocation of unit i 's margin

$$M^i = \sum_j (\beta_j^i g(S_j) - \alpha_j^i S_j) \Delta X_j. \quad (8.4)$$

Exercise. Use Equation 4.10 with Equation 4.8 and Equation 4.9 to derive a simple equation for M^i and then prove it is the equivalent of Equation 8.4.

Recall that the layer k cost of capital is

$$\iota_k := \frac{M_k}{Q_k} = \frac{P_k - S_k \Delta X_k}{\Delta X_k - P_k} = \frac{g(S_k) - S_k}{1 - g(S_k)}.$$

Here, we lean on the assumption of law invariance (see Section 4.4). We claim that for a law invariant pricing approach, the layer cost of capital must be the *same for all units*. Law invariance implies the risk measure is only concerned with the attachment probability of the layer and not with the cause of loss within the layer. If ι *within a layer* varied by unit,

then the allocation could not be law invariant. This crucial observation underlies the following logic.

Because cost of capital equals margin over capital, and seeing as both margin and cost of capital are known by layer and by unit, with the latter constant across units, we can compute unit capital by layer Q_k^i via

$$\iota_k := \frac{M_k}{Q_k} = \frac{M_k^i}{Q_k^i} \implies Q_k^i = \frac{M_k^i}{\iota_k}.$$

Substituting the details gives the following definition for the **NA of capital** in layer k to unit i :

$$Q_k^i = \frac{\beta_k^i g(S_k) - \alpha_k^i S_k}{g(S_k) - S_k} (1 - g(S_k)) \Delta X_k. \quad (8.5)$$

In words, taking Equation 8.5 and dividing numerator and denominator by capital in the layer shows

$$Q_k^i = \frac{\text{Margin for unit in layer}}{\text{Total margin in layer}} \times \text{Capital in layer} = \frac{\text{Margin for unit in layer}}{\text{Layer cost of capital}}.$$

The NA is the same as method 1 (divide by weighted average cost of capital) but *applied separately for each layer* of capital!

Exercise. Derive Equation 8.5.

Since $1 - g(S_k)$ is the proportion of of layer k 's limit devoted to capital, this says the proportion of capital allocated to unit i is given by the nicely symmetric expression

$$\frac{\beta_k^i g(S_k) - \alpha_k^i S_k}{g(S_k) - S_k}.$$

To determine total capital by unit, we sum across layers

$$Q^i := \sum_j Q_j^i.$$

Finally, we can determine the average cost of capital for unit i

$$\iota^i = \frac{M^i}{Q^i}.$$

This is illustrated for our example in Table 8.8, Table 8.9, and Table 8.10. In the last table, the columns are calculated as $M^i = \beta^i g(S) - \alpha^i S$, $\iota = (g - S)/(1 - g)$, $Q^i = M^i / \iota$, and the low rows M (respectively, Q) show sum products with ΔX .

Table 8.8. Calculation of loss shares by layer. The row L shows sum products with ΔX .

k	Unit A X^1	Unit B X^2	Unit C X^3	Total X	p	S	ΔX	$\alpha^1 S$	$\alpha^2 S$	$\alpha^3 S$
0	0	0	0	0	0	1	22	0.325	0.441	0.234
1	15	7	0	22	0.1	0.9	6	0.257	0.409	0.234
2	15	13	0	28	0.1	0.8	8	0.203	0.362	0.234
3	5	20	11	36	0.1	0.7	4	0.189	0.307	0.204
4	10	24	6	40	0.4	0.3	15	0.089	0.067	0.144
5	26	19	10	55	0.1	0.2	10	0.042	0.032	0.126
6	17	8	40	65	0.1	0.1	35	0.016	0.020	0.064
7	16	20	64	100	0.1	0	0	0	0	0
$L = E_p$	13.4	18.3	14.9	46.6				13.4	18.3	14.9

Table 8.9. Calculation of premium shares by layer. Row P shows sum products with ΔX .

k	Unit A X^1	Unit B X^2	Unit C X^3	Total X	q	gS	ΔX	$\beta^1 S$	$\beta^2 S$	$\beta^3 S$
0	0	0	0	0	0	1	22	0.295	0.409	0.296
1	15	7	0	22	0.052	0.948	6	0.260	0.392	0.296
2	15	13	0	28	0.066	0.882	8	0.225	0.362	0.296
3	5	20	11	36	0.075	0.807	4	0.214	0.320	0.273
4	10	24	6	40	0.379	0.428	15	0.119	0.093	0.216
5	26	19	10	55	0.119	0.309	10	0.063	0.051	0.194
6	17	8	40	65	0.135	0.174	35	0.028	0.035	0.111
7	16	20	64	100	0.174	0	0	0	0	0
$L = E_p$	13.4	18.3	14.9	46.6						
$P = E_p$	14.109	18.637	20.819	53.565				14.109	18.637	20.819
$M = P - L$	0.709	0.337	5.919	6.965						

Table 8.10. Allocation of capital.

k	Total X	ΔX	M^1	M^2	M^3	Q^1	Q^2	Q^3	ι	δ
0	0	22	-0.030	-0.032	0.061	0	0	0	infinite	1
1	22	6	0.003	-0.017	0.061	0.003	-0.018	0.067	0.917	0.478
2	28	8	0.021	-0.001	0.061	0.031	-0.001	0.089	0.693	0.409
3	36	4	0.025	0.013	0.069	0.045	0.024	0.125	0.555	0.357
4	40	15	0.030	0.026	0.072	0.134	0.115	0.323	0.224	0.183
5	55	10	0.021	0.019	0.069	0.133	0.121	0.437	0.158	0.136
6	65	35	0.012	0.015	0.047	0.132	0.165	0.529	0.089	0.082
7	100	0	0	0	0	0	0	0		
E_q			0.709	0.337	5.919	8.411	8.691	29.333		
ι						0.084	0.039	0.202	0.150	

Observe that $(\sum_i \iota^i Q^i)/Q$ is 15%, the weighted cost of capital over all units is 15%, as assumed.

Notice that the first layer $k = 0$ has undefined cost of capital ι ; since $S_0 = 1$, we have $g(S_0) = 1$, $P_0 = \Delta X_0$, and $M_0 = Q_0 = 0$. Nonetheless, we have nonzero unit margins M_0^i summing to zero.

The smallest portfolio loss equals 22, meaning the first layer is entirely funded by loss payments and requires no capital in aggregate. However, there are cross-subsidies between the units within the layer. Were the three units to purchase insurance from an insurance company with only \$22 of assets, it would benefit the most volatile line, because it would capture an outsized proportion of the available assets in the case it had a large loss. The margins reflect these cross-subsidies. However because the layer is fully funded by losses, there is no capital and hence no aggregate margin, which is why the allocated margins sum to zero.

Although the cost of capital *within* each layer is the same for all units, the margin, the proportion of limit that is capital, and the proportion of that capital attributable to each unit, all vary across layers. Therefore, *average cost of capital generally varies by unit*. Table 8.11 shows the difference in results between the natural Wang allocation and the CCoC approaches done previously. (The natural CCoC is equivalent to the industry standard approach.) The story is complex. While both non-catastrophe lines receive more margin and more capital under the Wang allocation, it is not proportional; their rates of return are not only lower than the 15% benchmark, but they are markedly different.

Table 8.11. Comparison of Wang SRM and CCoC natural capital allocations.

Distortion	Metric	Unit A	Unit B	Unit C	Portfolio
Wang, 0.343	Capital Q	8.411	8.691	29.333	46.435
	Margin M	0.709	0.337	5.919	6.965
	Return ι	8.4%	3.9%	20.2%	15.0%
CCoC, 0.150	Capital Q	4.874	5.288	36.273	46.435
	Margin M	0.339	0.222	6.404	6.965
	Return ι	7.0%	4.2%	17.7%	15.0%

For a typical distortion other than CCoC, in line with Table 1.1, we find:

- Lower layers of assets, below the expected losses, have a high cost of capital, but, offsetting this, they are mostly funded by premium and require little supporting investor capital, resulting in high leverage.
- Higher layers of assets have a lower cost of capital but higher capital content: they are funded by capital to a greater degree.

- Low volatility units by definition tend to have losses close to their expected value, regardless of the value of total losses, and so consume relatively more of the expensive, lower layer capital and a smaller proportion of cheaper, higher asset layers. Offsetting this, there is more capital in higher layers than lower ones.
- High volatility units tend to have a larger proportion of total losses when the portfolio total loss is large, and so they consume a greater proportion of cheaper, higher layer capital.

We see all of these phenomena in the example.

With price, quantity, and proportion all variable the overall effect of the NA of capital is hard to predict. Since price, amount, and use are all correlated, it is never adequate to assume that the unit average cost of capital is the portfolio average cost of capital. This is the fundamental flaw in the industry standard approach, which uses the same cost of capital for all capital layers (see Section 4.7). The SRM-based method outlined here produces varying cost of capital by unit in a defensible manner—given a choice of distortion function. How to make that choice was discussed in Chapter 6.

8.4. Comparisons with Bodoff's layer capital allocation

Readers may have guessed that the NA is closely related to Bodoff's layer capital allocation (Bodoff 2007). This section will explain how they are correct and show the connections, similarities, and differences between the two approaches. Bodoff's approach uses α^i , which describes how loss in each layer is shared between the units, and applies it to *all* assets in the layer. Since $\sum_i \alpha^i = 1$ we get a decomposition

$$a = \int_0^a 1 \, dx = \sum_i \int_0^a \alpha^i(x) \, dx.$$

Thus, $a^i = \int_0^a \alpha^i(x) \, dx$ is a reasonable allocation of assets. However, to determine a premium a^i needs to be split into amounts $P^i + Q^i$ funded by policyholder premium and investor capital. For example, we could fall back to CCoC, resulting in a layer premium

$$(\alpha^i(x)S(x) + \delta\alpha^i(x)(1 - S(x))) + \nu\delta\alpha^i(x)(1 - S(x)),$$

where the first parenthetic term is expected loss (Equation 8.2) plus a margin δ times the unfunded liability (Table 4.1, P3), and the second term is investor capital (Table 4.1, Q). This reproduces distortion CCoC pricing. Picking other distortions, which allow for the unfunded liability to be split differently than losses, produces a range of other answers. Bodoff's paper is very insightful, but it falls short by not fully describing a premium allocation: further assumptions are still needed to convert allocated assets into allocated premiums. Table 8.12 shows Bodoff's asset allocation in the last four right-hand columns.

Table 8.12. Bodoff's asset allocations, $\alpha^i \Delta X$.

k	Total X	p	S	ΔX	$\alpha^1 \Delta X$	$\alpha^2 \Delta X$	$\alpha^3 \Delta X$	$\sum \alpha^i \Delta X$
0	0	0	1	22	7.152	9.694	5.154	22
1	22	0.1	0.9	6	1.713	2.726	1.562	6
2	28	0.1	0.8	8	2.033	3.624	2.343	8
3	36	0.1	0.7	4	1.082	1.753	1.164	4
4	40	0.4	0.3	15	4.471	3.343	7.186	15.000
5	55	0.1	0.2	10	2.108	1.615	6.277	10.000
6	65	0.1	0.1	35	5.600	7.000	22.400	35.000
7	100	0.1	0	0	0	0	0	0
Total	46.6				24.159	29.756	46.086	100.0

9. Calculations with aggregate

Example calculations using Python and R.

This chapter reproduces many of the exhibits from the main text using Python and the `aggregate` package. It also describes how to extend to more realistic examples. Throughout, we have included only terse subsets of code so the reader can focus on the results and the output. Instructions for obtaining the full code are provided below.

9.1. `aggregate` and Python

`aggregate` is a Python package that can be used to build approximations to compound (aggregate) probability distributions quickly and accurately, and to solve insurance, risk management, and actuarial problems using realistic models that reflect underlying frequency and severity. It delivers the speed and accuracy of parametric distributions to situations that usually require simulation, making it as easy to work with an aggregate (compound) probability distribution as the lognormal. `aggregate` includes an expressive language called DecL to describe aggregate distributions.

The `aggregate` package is available on PyPi,⁵ the source code is on GitHub at <https://github.com/mynl/aggregate>, and there is extensive documentation.⁶ The `Aggregate` class and DecL language are described in Mildenhall (2024). There is also an extensive series of videos introducing various capabilities available on YouTube.⁷

9.2. `aggregate` and R

Python packages can be used in R via the `reticulate` library. Python and R code can also be mixed in RMarkdown (now Quarto) files. A short video⁸ explains how to use `aggregate` from R.

9.3. Reproducing the code examples

To reproduce the Python code examples, you must set up your environment and then install `aggregate` using

```
pip install aggregate
```

⁵ <https://pypi.org/project/aggregate/>.

⁶ <https://aggregate.readthedocs.io/en/latest/>.

⁷ <https://www.youtube.com/playlist?list=PLQQxycbewjqMDmw0hfZdB6Rzm60Qcq3Ao>.

⁸ <https://www.youtube.com/watch?v=jA9rVMHVqI0>.

The formatting of the examples relies on `greater_tables`, which can be installed in the same way. The installation process will also install several standard packages, such as `numpy` and `pandas`, that it depends on. If you get an error message about other missing packages, you can install them with `pip`. All the code in the monograph runs on aggregate version 0.24.0 and should run on any later version. If you have an older version installed, you can update it with `pip install -U aggregate`.

Once `aggregate` is installed, start by importing basic libraries. The next code block shows the relevant `aggregate` function. The class `GT` from `greater_tables` is used to format dataframes consistently.

```
from aggregate import (
    Portfolio,                # creates multi-unit portfolios
    make_ceder_netter,        # models reinsurance
    Distortion                 # creates a Distortion function
)
from greater_tables import GT # consistent table format
```

Online resources. Complete code samples can be extracted from the online version of the monograph, available at <https://casact.github.io/capital-modeling/>. Each code block has a small copy icon in the upper right-hand corner. The code blocks on each page can all be shown or hidden using the `</>Code` control at the top of the page. To download the original `qmd` file, use the View Source option under the same control. Alternatively, the entire Quarto (previously RMarkdown) source can be downloaded from GitHub at <https://github.com/casact/capital-modeling>.

9.4. Reproducing the discrete example

9.4.1. Input data and setup

The InsCo example running through this monograph is based on the 10 events shown in Table 9.1, which reproduces Table 2.1, and adds the mean and CV. For reference, here is the code to create the underlying dataframe.

```
loss_sample = pd.DataFrame(
    {
        'A': [ 5,  7, 15, 15, 13,  5, 15, 26, 17, 16],
        'B': [20, 33, 13,  7, 20, 27, 16, 19,  8, 20],
        'C': [11,  0,  0,  0,  7,  8,  9, 10, 40, 64],
    }
)
loss_sample['total'] = loss_sample.sum(axis=1)
loss_sample['p_total'] = 1/10
print(loss_sample)
```


	A	B	C	total	p_total
0	5	20	11	36	0.1
1	7	33	0	40	0.1
2	15	13	0	28	0.1
3	15	7	0	22	0.1
4	13	20	7	40	0.1
5	5	27	8	40	0.1
6	15	16	9	40	0.1
7	26	19	10	55	0.1
8	17	8	40	65	0.1
9	16	20	64	100	0.1

Table 9.1. The 10 equally likely simulations underlying the basic examples.

Index	A	B	C	Total
0	5	20	11	36
1	7	33	0	40
2	15	13	0	28
3	15	7	0	22
4	13	20	7	40
5	5	27	8	40
6	15	16	9	40
7	26	19	10	55
8	17	8	40	65
9	16	20	64	100
E[X]	13.4	18.3	14.9	46.6
CV	0.453	0.412	1.324	0.455

Now we create an `aggregate Portfolio` class instance based on the `loss_sample` simulation output. The next block shows how to do this, using the `Portfolio.create_from_sample` method. `TMA1` is a label, and `loss_sample` is the dataframe created previously. The other arguments specify working with unit-sized buckets `bs=1` appropriate for our integer losses and to use $256 = 2^8$ buckets, `log2=8`.

```
wport = Portfolio.create_from_sample('TMA1', loss_sample, bs=1, log2=8)
```

The next line of code displays Table 9.2, which shows summary output from the object `wport` by accessing its `describe` attribute. The summary includes the mean, CV, and skewness. The model and estimated (`Est`) columns are identical because we specified the distribution directly; no simulation or Fast Fourier Transform-based numerical methods are employed.

```
fGT(wport.describe.drop(columns=['Err E[X]', 'Err CV(X)']))
```

Table 9.2. Summary statistics for the base example.

Unit	X	$E[X]$	$EstE[X]$	$CV(X)$	$EstCV(X)$	$Skew(X)$	$EstSkew(X)$
A	Freq	1		0			
	Sev	13.4	13.4	0.453	0.453	0.281	0.281
	Agg	13.4	13.4	0.453	0.453	0.281	0.281
B	Freq	1		0			
	Sev	18.300	18.300	0.412	0.412	0.269	0.269
	Agg	18.300	18.300	0.412	0.412	0.269	0.269
C	Freq	1		0			
	Sev	14.9	14.900	1.324	1.324	1.603	1.603
	Agg	14.9	14.900	1.324	1.324	1.603	1.603
Total	Freq	3		0			
	Sev	15.533	15.533	0.827		1.884	
	Agg	46.6	46.600	0.471	0.471	1.177	1.177

Next, we need to calibrate distortion functions to achieve the desired pricing. The example uses full capitalization throughout (assets equal the maximum loss), which is equivalent to a 100% percentile level (see Table 9.5) and assumes the cost of capital is 15%. The code runs the calibration by calling `wport.calibrate_distortions` with arguments for the capital percentile level `Ps` and cost of capital `COCs`. The arguments are passed as lists because the routine can be used to solve for several percentile levels and costs simultaneously. Table 9.3 shows the resulting parameters, which match those in Table 4.5.

```
# calibrate to 15% return at p=100% capital standard
wport.calibrate_distortions(Ps=[1], COCs=[.15]);
```

Table 9.3. Distortion functions calibrated to 15% return on full capital.

Method	P	Parameter	Error
CCoC	53.565	0.150	0
PH	53.565	0.720	0.000
Wang	53.565	0.343	0.000
Dual	53.565	1.595	-0.000
TVaR	53.565	0.271	0.000

Figure 9.1 shows the resulting distortion g functions (top) and the probability adjustment $Z = q/p$ (bottom). The red star indicates a probability mass at $p = 1$.

For technical reasons (discussed in Section 4.3 as well as Section 14.1.5 in Milnehall and Major (2022), see especially Figure 14.2), we must summarize the 10 events by distinct totals. Routine code produces Table 9.4, reproducing Table 4.2.

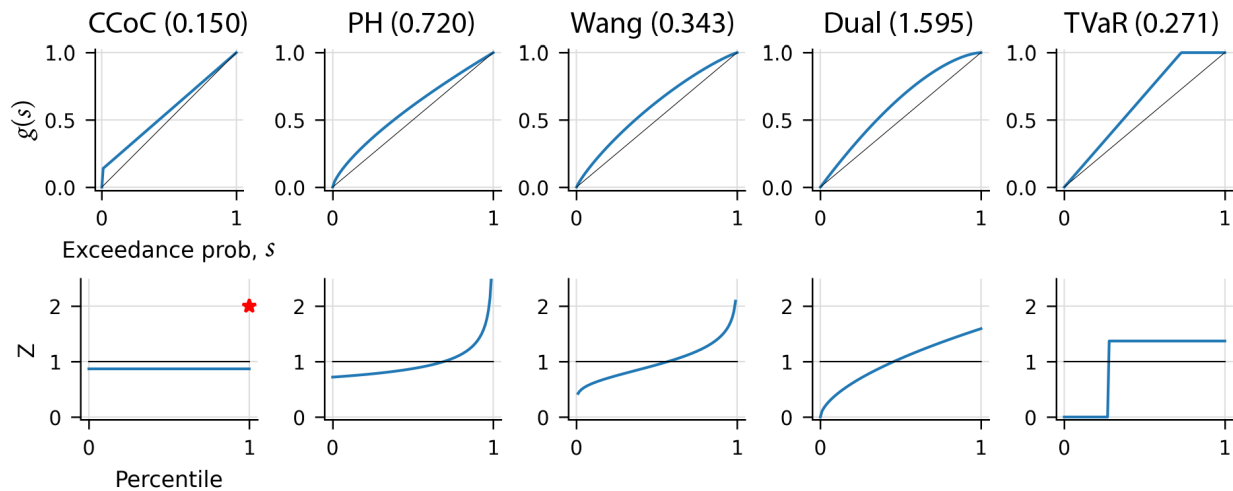


Figure 9.1. Distortion functions (top row) and distortion probability adjustment functions (bottom row).

Table 9.4. Assumptions for portfolio, summarized by distinct totals.

Index	A	B	C	p	Total
0	15	7	0	0.1	22
1	15	13	0	0.1	28
2	5	20	11	0.1	36
3	10	24	6	0.4	40
4	26	19	10	0.1	55
5	17	8	40	0.1	65
6	16	20	64	0.1	100
$E[X]$	13.4	18.3	14.9	1	46.6
Plan	13.9	18.7	19.6	1	52.2

The examples describe a VaR- or TVaR-based capital standard (see Table 3.1), but for this example it works out to be the same as a fully capitalized 100% standard, as shown by the values in Table 9.5. The `wport` methods `q` and `tvar` compute quantiles and TVaRs for each unit and the total.

Table 9.5. Quantiles (VaR) and TVaR at different probability levels.

p	Quantile	TVaR
0.8	55	82.500
0.850	65	88.333
0.9	65	100
1	100	100

9.4.2. Premiums by distortion

Table 4.11 shows premium assuming a CCoC distortion with a 15% return applied using a 100% capital standard. The premium P row shows what Mildenhall and Major (2022) call the linear NA premium. Table 9.6 reproduces this exhibit, confirming an equal 15% returns across all allocated capital. The code illustrates the power of `aggregate`. The first line creates a CCoC distortion using the `Distortion` class; note the argument is risk discount δ . The second line uses the `price` method to compute the linear allocation, which it returns in a structure with various diagnostic information, including Table 9.6.

```
ccoc = Distortion.ccoc(.15 / 1.15)
pricing_info = wport.price(1, ccoc, allocation='linear')
```

Table 9.6. Industry standard approach pricing using the CCoC distortion.

Index	A	B	C	Total
L	13.4	18.3	14.900	46.6
a	16.000	20.000	64.000	100
Q	2.261	1.478	42.696	46.435
P	13.739	18.522	21.304	53.565
M	0.339	0.222	6.404	6.965
COC	0.150	0.150	0.150	0.150

Table 4.3, Table 4.4, and Table 4.10 compute $\rho(X)$ for ρ the CCoC SRM, using the $\rho(X) = \int g(S(x)) dx$ (second table) and $\rho(X) = \int x g'(S(x)) f(x) dx$ (third) representations. The numbers needed for these calculations are shown in Table 9.7, where $q = g'(S(x)) f(x)$. These are extracted from the dataframe `wport.density_df`, which contains the probability mass, density, and survival functions for the portfolio, among other facts. The last row carries out the calculations and confirms the two methods give

Table 9.7. Industry standard approach pricing: raw ingredients and computed means.

Loss	p_{total}	F	S	gS	q	ΔX
22	0.1	0.1	0.9	0.913	0.087	6
28	0.1	0.2	0.8	0.826	0.087	8
36	0.1	0.3	0.7	0.739	0.087	4
40	0.4	0.7	0.3	0.391	0.348	15
55	0.1	0.8	0.2	0.304	0.087	10
65	0.1	0.9	0.1	0.217	0.087	35
100	0.1	1	0	0	0.217	0
$\rho(X)$				53.565	53.565	

the same result, the total under gS using the former representation and under q using the latter representation.

Table 9.8 shows the results for the other standard distortions. The calculations match those of the Wang in Table 9.8c. All of the calibrated distortions are carried in the dictionary `wport.dists`. These tables also include the NA of capital (a process we do not recommend, but that is described in Section 8.3). The `Q` row matches the calculation of allocated capital shown in Table 8.10.

Table 9.8. Pricing by distortion by unit.

(a) CCoC					(b) PH				
Statistic	A	B	C	Total	Statistic	A	B	C	Total
L	13.400	18.300	14.900	46.600	L	13.400	18.300	14.900	46.600
a	18.613	23.810	57.577	100.0	a	21.133	24.914	53.952	100.000
Q	4.874	5.288	36.273	46.435	Q	7.074	6.565	32.796	46.435
P	13.739	18.522	21.304	53.565	P	14.060	18.349	21.156	53.565
M	0.339	0.222	6.404	6.965	M	0.660	0.049	6.256	6.965
COC	0.070	0.042	0.177	0.150	COC	0.093	0.008	0.191	0.150

(c) Wang					(d) Dual				
Statistic	A	B	C	Total	Statistic	A	B	C	Total
L	13.400	18.300	14.900	46.600	L	13.400	18.300	14.900	46.600
a	22.520	27.328	50.152	100.000	a	22.999	28.260	48.741	100.0
Q	8.411	8.691	29.333	46.435	Q	8.873	9.143	28.419	46.435
P	14.109	18.637	20.819	53.565	P	14.127	19.117	20.322	53.565
M	0.709	0.337	5.919	6.965	M	0.727	0.817	5.422	6.965
COC	0.084	0.039	0.202	0.150	COC	0.082	0.089	0.191	0.150

(e) TVaR				
Statistic	A	B	C	Total
L	13.400	18.300	14.900	46.600
a	22.817	29.659	47.525	100.000
Q	9.034	9.247	28.154	46.435
P	13.783	20.412	19.371	53.565
M	0.383	2.112	4.471	6.965
COC	0.042	0.228	0.159	0.150

9.4.3. Reinsurance analysis

Chapter 5 analyzes a possible 35 xs 65 aggregate stop loss reinsurance contract. The setup to analyze this contract, using a separate `Portfolio` object, is shown next. The first two

lines apply reinsurance; `[1, 35, 65]` specifies 100% of the 35 xs 65 layer, and `c` and `n` are functions mapping gross to ceded and net. The last line creates a new `Portfolio` object from the net and ceded aggregate losses.

```
loss_sample_re = loss_sample.copy()
c, n = make_ceder_netter([(1, 35, 65)])
loss_sample_re['Net'] = n(loss_sample_re.total)
loss_sample_re['Ceded'] = c(loss_sample_re.total)
loss_sample_re = loss_sample_re[['Net', 'Ceded', 'p_total']]

# build Portfolio object with Net, Ceded units
wport_re = Portfolio.create_from_sample('WGS3',
    loss_sample_re[['Net', 'Ceded', 'p_total']], bs=1, log2=8)
```

Table 9.9 shows the standard summary statistics. It should be compared to Table 5.1. Table 9.10 shows the allocated pricing to the reinsurance and net across the five standard distortions; compare with the last row of Table 5.1 and discussion in Section 5.2.

Table 9.9. Summary statistics created by the reinsurance `Portfolio` object.

Unit	X	$E[X]$	$EstE[X]$	$CV(X)$	$EstCV(X)$	$Skew(X)$	$EstSkew(X)$
Net	Freq	1		0			
	Sev	43.1	43.100	0.317	0.317	0.369	0.369
	Agg	43.1	43.100	0.317	0.317	0.369	0.369
Ceded	Freq	1		0			
	Sev	3.5	3.500	3	3.000	2.667	2.667
	Agg	3.5	3.500	3	3.000	2.667	2.667
Total	Freq	2		0			
	Sev	23.3	23.300	0.998		0.340	
	Agg	46.6	46.600	0.370	0.370	0.788	0.788

9.5. A more realistic example

In this section, we create a series of exhibits analogous to those in Section 9.4 for an example with more realistic assumptions. It is included to show how `aggregate` can be used to solve real-world problems, hopefully motivating you to explore it further. The analysis steps are:

1. Create realistic by-unit frequency and severity distributions using Fast Fourier Transforms with independent units.
2. Sample the unit distributions with correlation induced by Iman-Conover (Section 8.2).
3. Build a `Portfolio` from the correlated sample.
4. Calibrate distortions and compute unit pricing for each distortion.
5. Apply by-unit, per-occurrence reinsurance and examine pricing impact.

Table 9.10. Pricing by distortion for ceded, net, and total (gross).

(a) CCoC				(b) PH			
Statistic	Ceded	Net	Total	Statistic	Ceded	Net	Total
L	3.500	43.100	46.600	L	3.500	43.100	46.600
a	30.013	69.987	100.0	a	23.383	76.617	100.000
Q	22.405	24.030	46.435	Q	16.721	29.714	46.435
P	7.609	45.957	53.565	P	6.662	46.903	53.565
M	4.109	2.857	6.965	M	3.162	3.803	6.965
COC	0.183	0.119	0.150	COC	0.189	0.128	0.150

(c) Wang				(d) Dual			
Statistic	Ceded	Net	Total	Statistic	Ceded	Net	Total
L	3.500	43.100	46.600	L	3.500	43.100	46.600
a	20.237	79.763	100.000	a	18.539	81.461	100.0
Q	14.150	32.284	46.435	Q	13.125	33.310	46.435
P	6.087	47.478	53.565	P	5.415	48.151	53.565
M	2.587	4.378	6.965	M	1.915	5.051	6.965
COC	0.183	0.136	0.150	COC	0.146	0.152	0.150

(e) TVaR			
Statistic	Ceded	Net	Total
L	3.500	43.100	46.600
a	17.679	82.321	100.000
Q	12.876	33.559	46.435
P	4.803	48.762	53.565
M	1.303	5.662	6.965
COC	0.101	0.169	0.150

The aggregate DecL programming language makes it easy to specify frequency-severity compound distributions. In the code chunk below, the four lines inside the `build` statement are the DecL program. The triple quotes are Python shorthand for entering a multiline string. The first program line, beginning `agg Auto`, creates a distribution with an expected loss of 5,000 (think: losses in 000s), severity from the 5,000 xs 0 layer of a lognormal variable with a mean of 50 and a CV of 2, and gamma mixed-Poisson frequency with a mixing parameter (CV) of 0.2. The other two lines are analogous. The `build` statement runs the DecL program and creates a `Portfolio` object with relevant compound distributions by unit. It also sums the unit distributions as though they were independent—we will introduce some correlation later.

```
from aggregate import build
port = build("""
```

```

port Units
agg Auto 5000 loss 5000 xs 0 sev lognorm 50 cv 2 mixed gamma 0.2
agg GL 3000 loss 2500 xs 0 sev lognorm 75 cv 3 mixed gamma 0.3
agg WC 7000 loss 25000 xs 0 sev lognorm 5 cv 10 mixed gamma 0.25
""")

```

Table 9.11 shows the by-unit frequency and severity statistics and the total statistics assuming the units are independent. The deviation between the Fast Fourier Transform-generated compound distributions and the requested specifications are negligible. Figure 9.2 plots the unit and total densities on a nominal and log scale. The effect of WC driving the tail, via thicker severity and higher occurrence limit, is clear on the log plot.

Table 9.11. Unit frequency, severity and compound assumptions, and portfolio total, showing requested and model achieved and key statistics.

Unit	X	$E[X]$	$EstE[X]$	$CV(X)$	$EstCV(X)$	$Skew(X)$	$EstSkew(X)$
Auto	Freq	100.0	0	0.224	0	0.402	0
	Sev	49.982	49.982	1.973	1.973	10.678	10.678
	Agg	5,000	5,000.0	0.298	0.298	0.699	0.699
GL	Freq	41.008	0	0.338	0	0.604	0
	Sev	73.156	73.156	2.377	2.378	7.385	7.385
	Agg	3,000	3,000.0	0.502	0.502	1.024	1.024
WC	Freq	1,401.1	0	0.251	0	0.500	0
	Sev	4.996	4.945	9.103	9.198	147.9	147.9
	Agg	7,000	7,000.0	0.350	0.350	1.776	1.775
Total	Freq	1,542.1	0	0.229	0	0.496	0
	Sev	9.727	9.681	6.123	0	68.886	0
	Agg	15,000	15,000.0	0.216	0.216	0.939	0.938

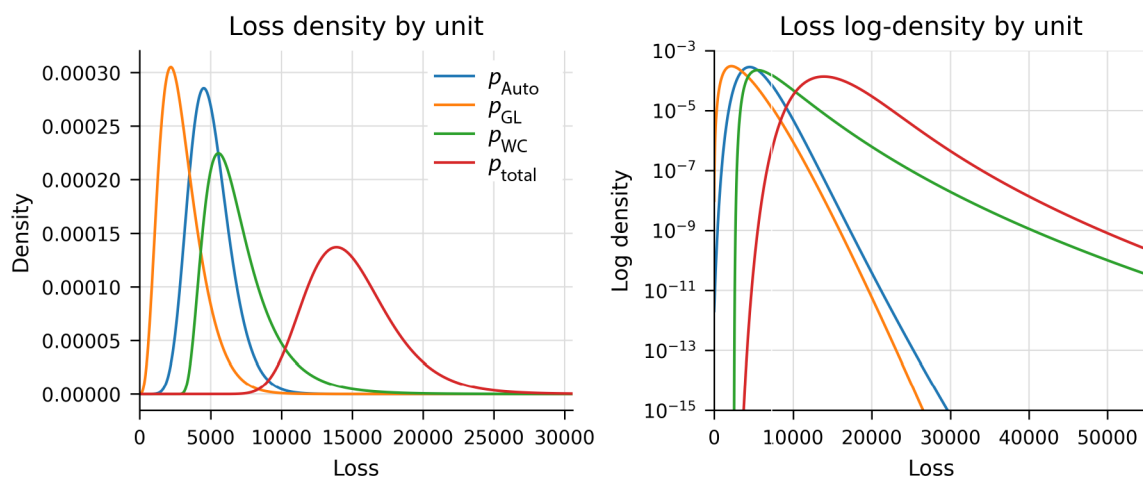


Figure 9.2. By-unit loss densities (left) and logdensities (right).

Next, we sample from the unit distributions and shuffle using Iman-Conover to achieve the desired correlation shown in Table 9.12. This matrix comes from a separate analysis. The revised statistics (note higher total CV), quantiles, and achieved correlation are shown in Table 9.13 and Table 9.14.

Table 9.12. Desired correlation matrix, as input to Iman-Conover.

Index	Auto	GL	WC
Auto	1	0.5	0.4
GL	0.5	1	0.1
WC	0.4	0.1	1

Table 9.13. Key statistics from sample with Iman-Conover induced correlation.

Index	Auto	GL	WC	Total
Count	1,000	1,000	1,000	1,000
Mean	4,996.5	2,989.2	7,017.9	15,003.6
SD	1,514.1	1,532.9	2,522.4	4,087.7
Min	1,118	133	3,148	6,576
25%	3,886.8	1,844.8	5,285.2	12,127.8
50%	4,884	2,691	6,459	14,573.5
75%	5,914.2	3,932.2	8,024.5	17,285
Max	11,957	10,184	22,877	34,617
CV	0.303	0.513	0.359	0.272

Table 9.14. Achieved between-unit correlation.

Index	Auto	GL	WC	Total
Auto	1	0.494	0.376	0.788
GL	0.480	1	0.070	0.601
WC	0.373	0.090	1	0.783
Total	0.788	0.625	0.732	1

Table 9.15 shows the output of using the correlated sample to build a `Portfolio` object. This uses the samples directly, proxying the aggregate loss as a compound with degenerate frequency distribution identically equal to one and severity equal to the desired distribution. Hence the frequency rows show expectation one with zero CV.

Table 9.16 shows the parameters for distortions calibrated to achieve an overall 15% return at a 99.5% capital standard. Figure 9.3 plots the distortions and probability adjustment functions; compare this with Figure 9.1. These distortions are slightly more risk averse (higher parameters except PH, driving higher indicated prices).

Table 9.15. Portfolio statistics reflecting the correlation achieved by the Iman-Conover sample.

Unit	X	$E[X]$	$EstE[X]$	$CV(X)$	$EstCV(X)$	$Skew(X)$	$EstSkew(X)$
Auto	Freq	1		0			
	Sev	4,996.5	4,996.5	0.303	0.303	0.625	0.625
	Agg	4,996.5	4,996.5	0.303	0.303	0.625	0.625
GL	Freq	1		0			
	Sev	2,989.2	2,989.2	0.513	0.513	0.905	0.905
	Agg	2,989.2	2,989.2	0.513	0.513	0.905	0.905
WC	Freq	1		0			
	Sev	7,017.9	7,017.9	0.359	0.359	1.766	1.766
	Agg	7,017.9	7,017.9	0.359	0.359	1.766	1.766
total	Freq	3		0			
	Sev	5,001.2	5,001.2	0.505		1.203	
	Agg	15,003.6	15,003.6	0.221	0.221	0.925	0.925

Table 9.16. Distortion parameters for target total return at the regulatory capital standard.

Method	P	Parameter	Error
CCoC	16,800.3	0.150	0
PH	16,800.3	0.654	0
Wang	16,800.3	0.437	0.000
Dual	16,800.3	1.756	-0.000
TVaR	16,800.3	0.299	0.000

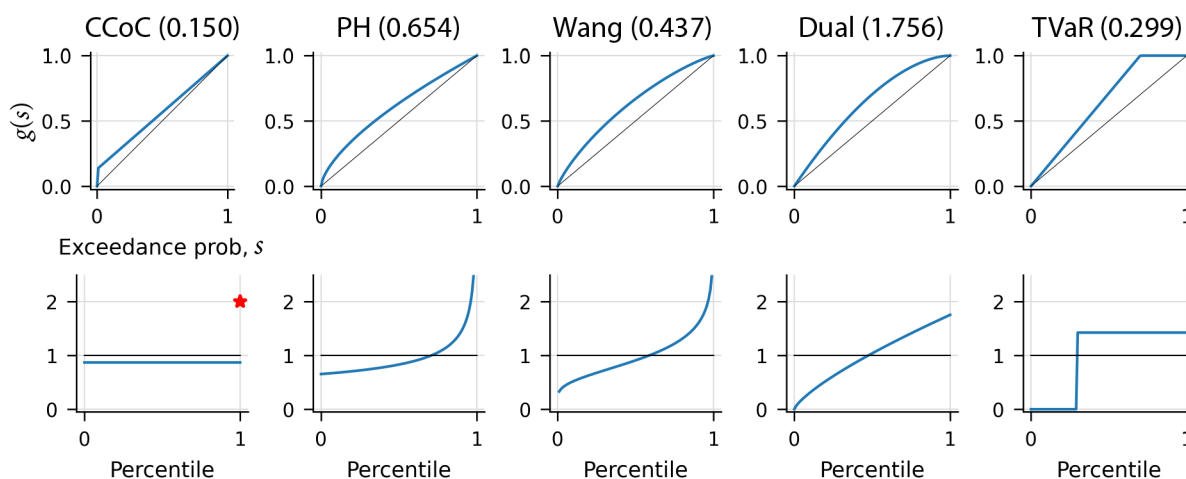
**Figure 9.3. Distortion functions and distortion probability adjustment functions.**

Table 9.17. Pricing by distortion.

(a) CCoC					(b) PH				
Statistic	Auto	GL	WC	Total	Statistic	Auto	GL	WC	Total
L	4,991.9	2,986.7	7,007.1	14,985.7	L	4,991.9	2,986.7	7,007.1	14,985.7
a	7,483.3	3,225.2	18,189.5	28,898.0	a	8,950.3	5,489.7	14,458.0	28,898.0
Q	2,222.6	269.8	9,605.3	12,097.7	Q	3,473.1	2,159.2	6,465.4	12,097.7
P	5,260.7	2,955.5	8,584.1	16,800.3	P	5,477.2	3,330.6	7,992.6	16,800.3
M	268.9	-31.227	1,577.0	1,814.7	M	485.3	343.9	985.5	1,814.7
COC	0.121	-0.116	0.164	0.150	COC	0.140	0.159	0.152	0.150

(c) Wang					(d) Dual				
Statistic	Auto	GL	WC	Total	Statistic	Auto	GL	WC	Total
L	4,991.9	2,986.7	7,007.1	14,985.7	L	4,991.9	2,986.7	7,007.1	14,985.7
a	9,171.0	5,651.9	14,075.0	28,898.0	a	9,264.8	5,753.9	13,879.3	28,898.0
Q	3,656.4	2,263.5	6,177.8	12,097.7	Q	3,719.5	2,319.4	6,058.8	12,097.7
P	5,514.7	3,388.4	7,897.3	16,800.3	P	5,545.3	3,434.5	7,820.5	16,800.3
M	522.8	401.7	890.2	1,814.7	M	553.4	447.8	813.4	1,814.7
COC	0.143	0.177	0.144	0.150	COC	0.149	0.193	0.134	0.150

(e) TVaR				
Statistic	Auto	GL	WC	Total
L	4,991.9	2,986.7	7,007.1	14,985.7
a	9,303.8	5,791.6	13,802.6	28,898.0
Q	3,727.5	2,321.0	6,049.3	12,097.7
P	5,576.3	3,470.7	7,753.3	16,800.3
M	584.5	484.0	746.2	1,814.7
COC	0.157	0.209	0.123	0.150

Table 9.17 compares the indicated pricing by unit and distortion. It is computed using exactly the same method described for the simple discrete example. The calculations are easy enough that they can be performed in a spreadsheet, but that is not recommended unless you are very patient. Table 9.18 shows just the implied loss ratios. The total loss ratios are all the same because the distortions are all calibrated to the same overall total premium. Since WC has the heaviest tail, it gets the lowest target loss ratio under CCoC, which is the most tail-risk averse. In fact, CCoC is so tail-risk averse that GL (which offers some diversification because of the correlation structure) gets a target over 100%. This is typical CCoC behavior—cat is really expensive; everything else is cheaper. The reason is clear when you remember CCoC pricing is just a weighted average of the mean and the max. The other distortions are more reasonable. TVaR and dual are quite similar and

Table 9.18. Implied loss ratio by unit, by pricing, and by distortion (extract from previous tables).

Method	Auto	GL	WC	Total
Dist CCoC	94.9%	101.1%	81.6%	0.892
Dist PH	91.1%	89.7%	87.7%	0.892
Dist Wang	90.5%	88.1%	88.7%	0.892
Dist Dual	90.0%	87.0%	89.6%	0.892
Dist TVaR	89.5%	86.1%	90.4%	0.892

focus more on volatility risk. Since all three lines are similarly volatile, it is no surprise their target loss ratios are also similar under this view.

This analysis—and indeed the whole monograph—has ignored several important elements: expenses, taxes, and the time value of money. Expenses are relatively easy to incorporate: acquisition expenses are typically amortized over the policy term, administration expenses span the policy term and claim payout period, and claim expenses can be bundled with losses. Taxes, however, are a Byzantine labyrinth requiring careful modeling at the level of legal entity and country. Their treatment is up to each insurer, and it's hard to offer even vague general guidance. Finally, time value of money brings real conceptual and computational difficulty (as opposed to the man-made complexity of taxes). It is the quantum gravity of insurance pricing. For some early thoughts, see the forthcoming monograph *Pricing Multi-Period Insurance Risk* by your intrepid second author.

Appendices

A.1. Table of symbols

Table A.1. Symbols used in this monograph.

Symbol	Interpretation	Reference
α_k^i	unit i share of layer k expected loss	Section 8.3
β_k^i	unit i share of layer k premium	Section 8.3
δ	rate of discount, $\delta = \iota / (1 + \iota)$	Section 4.1
ΔX_k	layer k limit, $\Delta X_k = X_{k+1} - X_k$	Section 4.4
ϵ	scaling parameter	Section 3.3, Section 5.5
ϵ_-	minimum scale	Section 5.5.2
ϵ_+	maximum scale	Section 5.5.2
ι	expected return, cost of capital, $\iota = M/Q$	Section 4.1
ι^i	unit i cost of capital	Section 8.3
ι_k	layer k cost of capital	Section 8.3
$\kappa^i(x)$	unit i conditional expected loss, $\kappa^i(x) := E[X^i X = x]$	Section 4.3
$\nabla^i \phi(X)$	partial derivative with respect to scaling	Section 3.3
ν	discount factor, $\nu = 1 / (1 + \iota)$	Section 4.1
$\Phi(z)$	standard Gaussian cumulative distribution function	Section 4.6
$\rho()$	pricing risk measure	Section 1.6
θ	weight parameter for bi-TVaR	Section 6.1
a	assets	Section 1.6
$a()$	capital risk measure	Section 1.6
a^i	allocation of assets to unit i (industry standard)	Section 4.7
$g(s)$	distortion function	Section 4.4
i	index to portfolio unit	Section 2.2
j	event index, $j = 1, \dots, n$	Section 2.2
k	layer index, $k = 0, \dots, n - 1$	Section 4.4
L	expected loss	Section 4.1
L^i	unit i expected loss	Section 4.7, Section 4.5

Symbol	Interpretation	Reference
M	margin, $P = L + M$	Equation 4.1, Section 4.1
M^i	unit i margin $M^i = P^i - L^i$	Section 4.5
m	number of units	Section 2.2
n	number of events	Section 2.2
p_j	event probability $p_j = P(X = X_j)$	Section 2.2
P	premium $P = \rho(X)$	Section 1.6
P_k	layer k premium, $P_k = g(S_k)\Delta X_k$	Section 4.4
P_P	plan premium	Section 5.5.2
P_R	required premium	Section 5.5.2
P^i	unit i premium	Section 4.7, Section 4.5
q_k	distorted probability, $q_k = g(S_{k-1}) - g(S_k)$	Section 4.4
Q	capital $Q = a - P$	Section 1.6
Q^i	unit i capital	Section 8.3
Q_k^i	unit i capital in layer k	Section 8.3
Q_P	plan capital, $Q_P = a - P_P$	Section 5.5.2
r	EVA/capital ratio, $r = V/Q_P$	Section 5.5.2
S_k	layer k attachment probability, $S_k = Pr(X > X_k)$	Section 4.4
$\text{TVaR}_{0.99}$	InsCo example capital standard	Section 3.3
V	EVA, economic value added, $V = P_P - P_R$	Section 5.5.2
$\text{VaR}_{0.85}$	example alternative risk measure	Section 3.3
X	portfolio liability random variable	Section 1.6
X_0	phantom portfolio loss $X_0 = 0$ with probability $p_0 = 0$	Section 4.3
X_j^i	loss to unit i in event j	Section 2.2
X_j	portfolio loss in event j	Section 2.2
X_k	layer payout threshold (attachment)	Section 4.4
$X \wedge a$	portfolio payout to policyholders	Section 1.6
$X^i(a)$	unit i capped losses	Section 4.3
X^i	loss to unit i (random variable)	Section 2.2
Z_k	event weight, likelihood ratio, $Z_k = q_k/p_k$	Section 4.4

A.2. Actuarial Standards of Practice

“The Actuarial Standards Board (ASB) sets standards for appropriate actuarial practice in the United States through the development and promulgation of Actuarial Standards of Practice (ASOPs). These ASOPs describe the procedures an actuary should follow when performing actuarial services and identify what the actuary should disclose when communicating the results of those services.”⁹ ASOPs can be found at <https://www.actuarialstandardsboard.org/standards-of-practice/>.

Table A.2. ASOPs that are relevant to capital modeling.

Number	Title	Reference
7	Analysis of Life, Health, or Property/Casualty Insurer Cash Flows	Chapter 2
23	Data Quality	Chapter 2
30	Treatment of Profit and Contingency Provisions and the Cost of Capital in Property/Casualty Insurance Ratemaking	Chapter 4
46	Risk Evaluation in Enterprise Risk Management	Chapter 2, Chapter 3
47	Risk Treatment in Enterprise Risk Management	Section 1.3, Section 3.3
55	Capital Adequacy Assessment	Chapter 3
56	Modeling	Chapter 2
58	Enterprise Risk Management	Chapter 2, Chapter 3

The ASB repealed ASOP Nos. 46 and 47 in December 2024 and replaced them with ASOP No. 58, Enterprise Risk Management, to reflect the developments since 2012, to better reflect today’s ERM practices and terminology, and to align with ASOP No. 55.

Table A.3. ASOPs that are potentially relevant to capital modeling.

Number	Title
12	Risk Classification (for All Practice Areas)
13	Trending Procedures in Property/Casualty Insurance
19	Appraisals of Casualty, Health, and Life Insurance Businesses
20	Discounting of Property/Casualty Claim Estimates
25	Credibility Procedures
29	Expense Provisions for Prospective Property/Casualty Risk Transfer and Risk Retention
36	Statements of Actuarial Opinion Regarding Property/Casualty Loss, Loss Adjustment Expense, or Other Reserves
38	Catastrophe Modeling (for All Practice Areas)
39	Treatment of Catastrophe Losses in Property/Casualty Insurance Ratemaking
41	Actuarial Communications
43	Property/Casualty Unpaid Claim Estimates
53	Estimating Future Costs for Prospective Property/Casualty Risk Transfer and Risk Retention

⁹ <https://www.actuarialstandardsboard.org/about-asb/>.

A.3. Glossary

Term	Description	Reference
Accounting	module that produces financial statements	Section 1.5
Allocation	module that distributes portfolio premium to units	Section 1.5
BCAR	Best's Capital Adequacy Ratio	Section 1.5
bi-TVaR	convex combination of TVaRs	Chapter 6
Business Operations	module that deals with loss experience	Section 1.5
capital	owner-provided funds	Section 1.1
Capital Adequacy	module dealing with risk and capital sufficiency	Section 1.5
capital risk measure	rule relating liabilities to required assets	Section 1.6
capital structure	mix of liabilities funding total assets	Section 1.5, Section 2.5
charter	formal document setting out the rationale	Section 1.2
coTVaR	conditional TVaR, Natural Allocation of TVaR	Section 3.3
comonotonic	random variables X and Y are comonotonic if they are nondecreasing functions of a third r.v.	Chapter 4
comonotonic additive	property of a risk measure ρ : if X and Y are comonotonic, then $\rho(X + Y) = \rho(X) + \rho(Y)$	Chapter 4
constant cost of capital (CCoC)	assumption that all layers require the same return	Section 4.4
distortion function	translates attachment probability to rate on line	Section 4.4
distorted expectation	expectation calculated with distorted probabilities q	Section 4.5, Section 4.8
distorted expected loss share	a unit's share of a layer's distorted expected loss, β_k^i	Section 8.3
distorted probability	after applying a distortion function, q replaces p	Section 4.4
Economic Scenario Generator	module to generate socioeconomic outcomes	Section 1.5, Section 2.4
economic value added (EVA)	profits beyond what is necessary for investors	Section 4.1

Term	Description	Reference
equal priority	rule for distributing funds among claimants	Section 4.3
expected loss share	a unit's share of a layer expected loss, α_k^i	Section 8.3
funding equation	states that assets are the sum of premiums and investor capital, $a = P + Q$	Equation 1.1, Section 1.6
gradient	derivative with respect to several variables	Section 3.3
InsCo	hypothetical insurer used as a simple example	Section 1.6
insureds	customers of InsCo, who pay premiums	Section 1.6
investors	owners of InsCo, who supply capital	Section 1.6
law invariance	risk measure depends only on distribution	Section 4.4
layer	segment of assets between two levels of portfolio loss	Section 4.4
layer funding equation	states that layer assets are the sum of premiums and investor capital, $\Delta X_k = P_k + Q_k$	Equation 4.4, Section 4.4
likelihood ratio	ratio of distorted to original probabilities	Section 4.4
margin	difference between premium and expected loss, $M = P - L$	Equation 4.1, Section 4.1
marginal approach	examine the effect of a small change; look at first derivative	Section 3.3
model	simplified representation of reality, usually mathematical	Section 1.1, Section 1.5
Natural Allocation (NA)	decomposition of portfolio premium into distorted expectation of unit losses	Section 4.5
Natural Allocation of capital	allocation of capital consistent with allocation of premium, Q_k^i	Section 8.3
normal copula	multivariate uniform distribution based on the normal	Section 8.2.5
objectivity	synonym for law invariance	Section 4.4
parameter uncertainty	lack of knowledge about parameters of a process, sometimes modeled as a distribution	Section 8.1
Pricing	module that assigns technical premium to the portfolio	Section 1.5
Pricing & Allocation	module that assigns technical premiums to units	Section 1.5
pricing risk measure	rule that relates liabilities to technical premiums	Section 1.6

Term	Description	Reference
pro forma	hypothetical financial statement	Section 1.5, Section 2.7
proxy model	simpler model of another model	Section 5.5
probabilistic database	stored sample of random outcomes	Section 2.2
process risk	randomness in outcomes from a well-defined and parameterized stochastic process	Section 8.1
Reinsurance	module that deals with risk transfer	Section 1.5
risk appetite	compensation/price limit for taking on risk	Section 1.5
risk tolerance	limit of risk the firm is willing to take on	Section 1.5
RVaR	Range Value at Risk, window version of VaR	Section 3.3
quantities of interest (QoI)	target metric, statistic, etc. of a simulation exercise	Section 8.1
Spectral Risk Measure (SRM)	risk measure determined by a distortion function	Section 4.4
state price	a contract that pays \$1 in a particular state of the world	Section 4.8
Tail Value at Risk (TVaR)	average of losses in a specified tail probability; also CVaR, TCE	Section 3.2
tranche	synonym for layer	Section 4.4
unit	portion of portfolio, e.g., line of business	Section 1.5
Value at Risk (VaR)	risk measure, quantile of a distribution	Section 1.5
XTVaR	excess TVaR, TVaR minus expected losses	Section 4.7

References

- Aas, Kjersti, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. 2009. "Pair-copula Constructions of Multiple Dependence." *Insurance: Mathematics and Economics* 44 (2): 182–98. <https://doi.org/10.1016/j.insmatheco.2007.02.001>.
- Actuarial Standards Board. 2007. "Property/Casualty Unpaid Claim Estimates." *Actuarial Standard of Practice No. 43*.
- Actuarial Standards Board. 2011. "Treatment of Profit and Contingency Provisions and the Cost of Capital in Property/Casualty Insurance Ratemaking." *Actuarial Standard of Practice No. 30*.
- Actuarial Standards Board. 2019. "Modeling." *Actuarial Standard of Practice No. 56*.
- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. "Coherent Measures of Risk." *Mathematical Finance* 9 (3): 203–28. <http://onlinelibrary.wiley.com/doi/10.1111/1467-9965.00068/abstract>.
- Blanchard, Ralph S., III. 2008. *Basic Insurance Accounting—Selected Topics*. July. CAS Study Note.
- Bodoff, Neil M. 2007. "Capital Allocation by Percentile Layer." *Variance* 3 (1): 13–30.
- Bolstad, William M., and James M. Curran. 2016. *Introduction to Bayesian Statistics*. John Wiley & Sons.
- Brehm, Paul J., Spencer M. Gluck, Kreps Rodney E., et al. 2007. *Enterprise Risk Analysis for Property & Liability Insurance Companies*. Guy Carpenter & Company, LLC.
- Bühlmann, Hans. 1985. "Premium Calculation from Top Down." *ASTIN Bulletin* 15 (2): 89–101. <https://doi.org/10.2143/ast.15.2.2015021>.
- Caramagno, Nicholas, David Mamane, and Liam Neilson. 2021. *An Introduction to IFRS 17 for P&C Actuaries*. Casualty Actuarial Society. https://www.casact.org/sites/default/files/2021-09/6I_IFRS_Study_Note.pdf.
- Casualty Actuarial Society. 2022. *Complete Text References for Exam 7, Syllabus of Basic Education*. https://www.casact.org/sites/default/files/2021-03/7_individual_text_ref.pdf.
- Chan, Cynthia, et al. 2024. *Insurance Rating Criteria*. Fitch Ratings. <https://www.fitchratings.com/research/insurance/insurance-rating-criteria-04-03-2024>.

- Cherny, Alexander, and Dilip Madan. 2009. "New Measures for Performance Evaluation." *Review of Financial Studies* 22 (7): 2571–606. <https://doi.org/10.1093/rfs/hhn081>.
- Conning. 2020. "A User's Guide to Economic Scenario Generation in Property/Casualty Insurance." *CAS Research Papers*. Casualty Actuarial Society. <https://www.casact.org/sites/default/files/2021-02/economic-scenario-generation-conning1020.pdf>.
- Cont, Rama, Romain Deguest, and Giacomo Scandolo. 2010. "Robustness and Sensitivity Analysis of Risk Measurement Procedures." *Quantitative Finance* 10 (6): 593–606. <https://doi.org/10.1080/14697681003685597>.
- Cummins, J. David. 1990. "Multi-Period Discounted Cash Flow Rate-Making Models in Property-Liability Insurance." *Journal of Risk and Insurance* 57 (1): 79–109.
- Cummins, J. David. 2000. "Allocation of Capital in the Insurance Industry." *Risk Management and Insurance Review*, 3 (1): 7–28. <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6296.2000.tb00013.x/full>.
- Daykin, Chris D., Teivo Pentikainen, and Martti Pesonen. 1993. *Practical Risk Theory for Actuaries*. Chapman & Hall/CRC Press.
- Delbaen, Freddy. 2000. "Coherent Risk Measures." *Blätter der DGVFM* 24 (4): 733–39. <https://doi.org/10.1007/BF02809088>.
- EIOPA. 2009. *Directive 138/2009/EC (Solvency II Directive)*. European Insurance and Occupational Pension Authority. https://www.eiopa.europa.eu/rulebook/solvency-ii-single-rulebook/directive-1382009ec-solvency-ii-directive_en.
- EIOPA. 2015. *Delegated Regulation (EU) 2015/35*. European Insurance and Occupational Pension Authority. https://www.eiopa.europa.eu/rulebook/solvency-ii-single-rulebook/delegated-regulation-eu-201535_en.
- Eling, Martin, and David Holzmüller. 2008. "An Overview and Comparison of Risk-Based Capital Standards." *Journal of Insurance Regulation* 26 (4): 31–60.
- Embrechts, Paul, and Giovanni Puccetti. 2010. "Risk Aggregation." In *Copula Theory and Its Applications*, 111–26. Springer. https://link.springer.com/chapter/10.1007/978-3-642-12465-5_5.
- Emma, Charles C., et al. 2000. *Dynamic Financial Models of Property-Casualty Insurers*. Casualty Actuarial Society. https://www.casact.org/sites/default/files/database/forum_00wforum_00wf317.pdf.
- Fairley, William B. 1979. "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results." *The Bell Journal of Economics* 10 (1): 192–210. <https://doi.org/10.2307/3003326>.

- Floreani, Alberto. 2011. "Risk Margin Estimation through the Cost of Capital Approach: Some Conceptual Issues." *Geneva Papers on Risk and Insurance: Issues and Practice* 36 (2): 226–53. <https://doi.org/10.1057/gpp.2011.2>.
- Friedland, Jacqueline. 2010. *Estimating Unpaid Claims Using Basic Techniques*. Casualty Actuarial Society. https://www.casact.org/sites/default/files/database/study_notes_friedland_estimating.pdf.
- Friedman, J., T. Hastie, and R. Tibshirani. 2008. *The Elements of Statistical Learning*. Second edition. Springer. <https://doi.org/10.1007/978-0-387-84858-7>.
- Gabaix, Xavier, and David I. Laibson. 2008. *The Seven Properties of Good Models*. Oxford University Press. <https://laibson.scholars.harvard.edu/publications/seven-properties-good-models>.
- Geisser, Seymour. 2017. *Predictive Inference*. Chapman & Hall/CRC Press.
- Golub, Gene H., and Charles F. Van Loan. 2013. *Matrix Computations*. Johns Hopkins University Press.
- Hansen, Lars Peter, and Thomas J. Sargent. 2008. "Robustness." In *Sampled-Data Models for Linear and Nonlinear Systems*. Springer. https://doi.org/10.1007/978-1-4471-5562-1_7.
- Heynderickx, W., J. Cariboni, W. Schoutens, and B. Smits. 2016. "The Relationship between Risk-Neutral and Actual Default Probabilities: The Credit Risk Premium." *Applied Economics* 48 (42): 4066–81. <https://doi.org/10.1080/00036846.2016.1150953>.
- IAA Risk Margin Working Group. 2009. *Measurement of Liabilities for Insurance Contracts: Current Estimates and Risk Margins*. International Actuarial Association.
- IAIS. 2008. *Standard on the Use of Internal Models for Regulatory Capital Purposes*. https://www.iaisweb.org/uploads/2022/01/16__Standard_No__2_2_7_on_the_use_of_Internal_Models_for_regulatory_capital_purposes.pdf.
- Iman, Ronald L., and William J. Conover. 1979. "The Use of the Rank Transform in Regression." *Technometrics* 21 (4): 499–509. <https://doi.org/10.1080/00401706.1979.10489820>.
- Jakobsen, Mathilde, Timothy Prince, Brad Herman, and Mahesh Mistry. 2020. *Understanding Universal BCAR*. A. M. Best.
- Jewson, Stephen, Trevor Sweeting, and Lynne Jewson. 2025. "Reducing Reliability Bias in Assessments of Extreme Weather Risk Using Calibrating Priors." *Advances in Statistical Climatology, Meteorology and Oceanography* 11 (1): 1–22. <https://doi.org/10.5194/asmo-11-1-2025>.
- Jouini, E., W. Schachermayer, and N. Touzi. 2008. "Optimal Risk Sharing for Law Invariant Monetary Utility Functions." *Mathematical Finance* 18 (2): 269–92. <https://doi.org/10.1111/j.1467-9965.2007.00332.x>.

- Karakuyu, Ali, Mark Button, et al. 2023. *Criteria | Insurance | General: Insurer Risk-Based Capital Adequacy—Methodology and Assumptions*. S&P Global. <https://www.spglobal.com/ratings/en/regulatory/article/231115-criteria-insurance-general-insurer-risk-based-capital-adequacy-methodology-and-assumptions-s12862217>.
- Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2019. *Loss Models: From Data to Decisions*. Fifth edition. John Wiley & Sons.
- Koca, Stephen, et al. 2023. *Statements of Actuarial Opinion on Property and Casualty Loss Reserves*. American Academy of Actuaries. <https://www.actuary.org/sites/default/files/2023-12/casualty-practice-note-saopclossreserves.pdf>.
- Koenker, Roger, and Kevin F. Hallock. 2001. “Quantile Regression.” *Journal of Economic Perspectives* 15 (4): 143–56. <https://www.aeaweb.org/articles?id=10.1257/jep.15.4.143>.
- Kovvali, Narayan. 2022. *Theory and Applications of Gaussian Quadrature Methods*. Springer Nature.
- Kozik, Thomas J. 1994. “Underwriting Betas—The Shadows of Ghosts.” *Proceedings of the Casualty Actuarial Society* 81 (154, 155): 303–29.
- Kreps, Rodney E. 1997. “Parameter Uncertainty in (Log)normal Distributions.” *Proceedings of the Casualty Actuarial Society* 84: 553–80. https://www.casact.org/sites/default/files/database/proceed_proceed97_97553.pdf.
- Major, John A. 1999. “Taking Uncertainty Into Account: Bias Issues Arising from Parameter Uncertainty in Risk Models.” *CAS Forum* Summer: 153–96.
- Major, John A., Ruodu Wang, and Micah G. Woolstenhulme. 2015. “The Most Dangerous Model: A Natural Benchmark for Assessing Model Risk.” *Society of Actuaries Monograph: Enterprise Risk Management Symposium*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611806.
- McClenahan, Charles L. 1999. “Insurance Profitability.” In *Actuarial Considerations Regarding Risk and Return*, 113–24. https://www.casact.org/sites/default/files/database/vfac_1999may.pdf.
- McNeil, Alexander J., Paul Embrechts, and Rüdiger Frey. 2005. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press.
- Meyers, Glenn G. 2019. *Stochastic Loss Reserving Using Bayesian MCMC Models*. Second edition. CAS Monograph Series. Casualty Actuarial Society.
- Mildenhall, Stephen J. 2004. “A Note on the Myers and Read Capital Allocation Formula.” *North American Actuarial Journal* 8 (2): 32–44. <https://www.tandfonline.com/doi/abs/10.1080/10920277.2004.10596135>.
- Mildenhall, Stephen J. 2017. “In Praise of Value at Risk.” *Actuarial Review* 44 (6): 48–

50. <https://ar.casact.org/in-praise-of-value-at-risk/>.
- Mildenhall, Stephen J. 2022. *When Is Premium Riskier Than Loss?* https://www.convexrisk.com/static/talks/MILDENHALL_When_Is_Premium_Riskier_Than_Loss_2022.pdf.

- Mildenhall, Stephen J. 2023a. *aggregate Documentation*. <https://aggregate.readthedocs.io/en/latest/>.
- Mildenhall, Stephen J. 2023b. “aggregate: Frequency-Severity Distributions Made Easy.” *Actuarial Review* 50 (3): 31–34. <https://ar.casact.org/aggregate-frequency-severity-distributions-made-easy/>.
- Mildenhall, Stephen J. 2024. “Aggregate: Fast, Accurate, and Flexible Approximation of Compound Probability Distributions.” *Annals of Actuarial Science* 19 (2): 1–40. <https://doi.org/10.1017/S1748499524000216>.
- Mildenhall, Stephen J. 2005. “Correlation and Aggregate Loss Distributions with an Emphasis on the Iman-Conover Method.” *Casualty Actuarial Society Forum* Winter: 103–204. https://www.casact.org/sites/default/files/database/forum_06wforum_06w107.pdf.
- Mildenhall, Stephen J., and John A. Major. 2022. *Pricing Insurance Risk: Theory and Practice*. John Wiley & Sons, Inc.
- Myers, Stewart C., and Richard A. Cohn. 1987. “A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation.” In *Fair Rate of Return in Property-Liability Insurance*. Springer.
- NAIC. 2024. *Risk-Based Capital*. National Association of Insurance Commissioners. <https://content.naic.org/insurance-topics/risk-based-capital>.
- Odomirok, Kathleen C., Liam M. Mcfarlane, Gareth L. Kennedy, and Justin J. Brenden. 2020. *Casualty Actuarial Society Financial Reporting through the Lens of a Property/Casualty Actuary*. Casualty Actuarial Society.
- Postek, Krzysztof, Alessandro Zocca, Joaquim Gromicho, and Jeffrey Kantor. 2024a. “Companion Jupyter Book for ‘Hands-On Mathematical Optimization with Python’.” The MO Book Group. <https://mobook.github.io/MO-book/intro.html>.
- Postek, Krzysztof, Alessandro Zocca, Joaquim Gromicho, and Jeffrey Kantor. 2024b. *Hands-On Mathematical Optimization with Python*. Cambridge University Press.
- Press, William H., Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 1992. *Numerical Recipes in C*. Second edition. Cambridge University Press.
- Reitano, Robert R. 1997. “Two Paradigms for the Market Value of Liabilities.” *North American Actuarial Journal* 1 (4): 104–22. <https://doi.org/10.1080/10920277.1997.10595657>.
- Robbin, Ira. 1992. *The Underwriting Profit Provision*. Casualty Actuarial Society Study Note.
- Smith, Ralph C. 2024. *Uncertainty Quantification: Theory, Implementation, and Applications*. SIAM.

- Society of Actuaries. 2024. *Economic Scenario Generators*. <https://www.soa.org/resources/tables-calcs-tools/research-scenario/>.
- Szkoda, Susan T., et al. 1995. *CAS Dynamic Financial Analysis Property/Casualty Insurance Companies Handbook*. Casualty Actuarial Society. https://www.casact.org/sites/default/files/database/forum_96wforum_96wf001.pdf.
- Tasche, Dirk. 1999. “Risk Contributions and Performance Measurement.” Report of the Lehrstuhl Fur Mathematische Statistik, TU Munchen.
- Venter, Gary G. 1991. “Premium Calculation Implications of Reinsurance without Arbitrage.” *ASTIN Bulletin* 21 (2): 223–30. <https://doi.org/10.2143/ast.21.2.2005365>.
- Venter, Gary G., John A. Major, and Rodney E. Kreps. 2006. “Marginal Decomposition of Risk Measures.” *ASTIN Bulletin* 36 (2): 375–413. <https://doi.org/10.2143/AST.36.2.2017927>.
- Wand, Matt P., and M. Chris Jones. 1994. *Kernel Smoothing*. CRC Press.
- Wang, Shaun S. 1996. “Premium Calculation by Transforming the Layer Premium Density.” *ASTIN Bulletin* 26 (1): 71–92. <https://doi.org/10.2143/AST.26.1.563234>.
- Wang, Shaun S. 1998. “Aggregation of Correlated Risk Portfolios: Models and Algorithms.” *Proceedings of the Casualty Actuarial Society* 86: 848–939. https://www.casact.org/sites/default/files/database/proceed_proceed98_980848.pdf.
- Wang, Shaun S. 2000. “A Class of Distortion Operators for Pricing Financial and Insurance Risks.” *Journal of Risk and Insurance* 67 (1): 15–36.
- Wüthrich, Mario V., and Michael Merz. 2015. *Stochastic Claims Reserving Manual: Advances in Dynamic Modeling*. Swiss Finance Institute Research Paper No. 15-34.
- Wüthrich, Mario V., Hans Bühlmann, Hansjörg Furrer, et al. 2010. *Market-Consistent Actuarial Valuation*. Volume 2. Springer.

ABOUT THE SERIES:

CAS monographs are authoritative, peer-reviewed, in-depth works focusing on important topics within property and casualty actuarial practice. For more information on the CAS Monograph Series, visit the CAS website at www.casact.org.



**Expertise. Insight.
Solutions.**

www.casact.org