Aggregating Risk Capital:
Lessons Learned From Basel II

Paul Embrechts

Department of Mathematics

and

RiskLab, ETH Zurich

www.math.ethz.ch/∼embrechts

Basel II: the basics

- late nineties: more refined (granular/risk sensitive) measurement of credit risk
- A new risk class: operational risk
- overall risk capital before/after equal
- three pillar approach

- McDonough (before Cooke) Ratio:

$$\frac{\text{total amount of capital}}{\text{risk weighted assets}} \geq 8\%$$

  - regulatory capital: Tiers 1, 2 and 3
  - risk weighted assets: that's where the Pillar 1 action is

- several quantitative impact studies

- start: 2007+?

- $X_1, \dots, X_d$ $\quad d$ 1-period risks, $\mathbf{X} = (X_1, \dots, X_d)$

- $\Psi\,(X_1, \dots, X_d)$ $\quad$ a financial "instrument" / "position"

  - $\Psi(\mathbf{x}) = s_d = \sum_{k=1}^{d} x_k$

  - $\Psi(\mathbf{x}) = m_d = \max\{x_1, \dots, x_d\}$

  - $\Psi(\mathbf{x}) = \sum_{k=1}^{d} (x_k - c_k)^+$ $\quad$ excess–of–loss

  - $\Psi(\mathbf{x}) = \left(\sum_{k=1}^{d} x_k - c\right)^+$ $\quad$ stop–loss

  - $\Psi(\mathbf{x}) = m_d I_{\{s_d > q_\alpha\}}$ $\quad$ digital trigger

  - ... (credit derivatives)

- a risk or pricing measure $\mathcal{R}$
  - $\mathcal{R}(Y) = \text{VaR}_\alpha(Y)$ Value–at–Risk
  - $\mathcal{R}(Y) = E\,(Y \mid Y > \text{VaR}_\alpha(Y))$ Expected Shortfall (warning)
  - $\mathcal{R}(Y) = E\left((Y-\mu)^k\right)$ moments
  - $\mathcal{R}(Y) = F_Y$, the distribution function
  - $\ldots$

- aim, calculate $\mathcal{R}(\Psi(\mathbf{X}))$ given
  - some information on $X_1, \ldots, X_d$ $(F_1, \ldots, F_d)$
  - for a specific $\mathcal{R}$

- Questions: is this a well–posed problem?
  is this relevant for practice?
  give examples!

We start with an example:

Operational Risk: The risk of losses resulting from inadequate or failed internal processes, people and systems, or external events. Included is legal risk, excluded are strategic/business risk and reputational risk

Examples: many!
Barings, ...

Recent QISs:

- Basel Committee QIS (Moscadelli 2004, 40 000+ observations)

- Federal Reserve Bank of Boston Loss–Date Collection Exercise (Dutta and Perry 2006, 50 000+ observations)

Similar issues (as in Basel II): Solvency 2 for Insurance

The OpRisk pillar 1 issue: Loss Distribution Approach (LDA)

- Data matrix $\mathcal{X}^t = \left\{X_{ij}^t : i = 1, \ldots, 8\,\text{BLs}, \quad j = 1, \ldots, 7\,\text{RTs}\right\}$ for year t, $t = 1, \ldots, T$ (only few years so far)

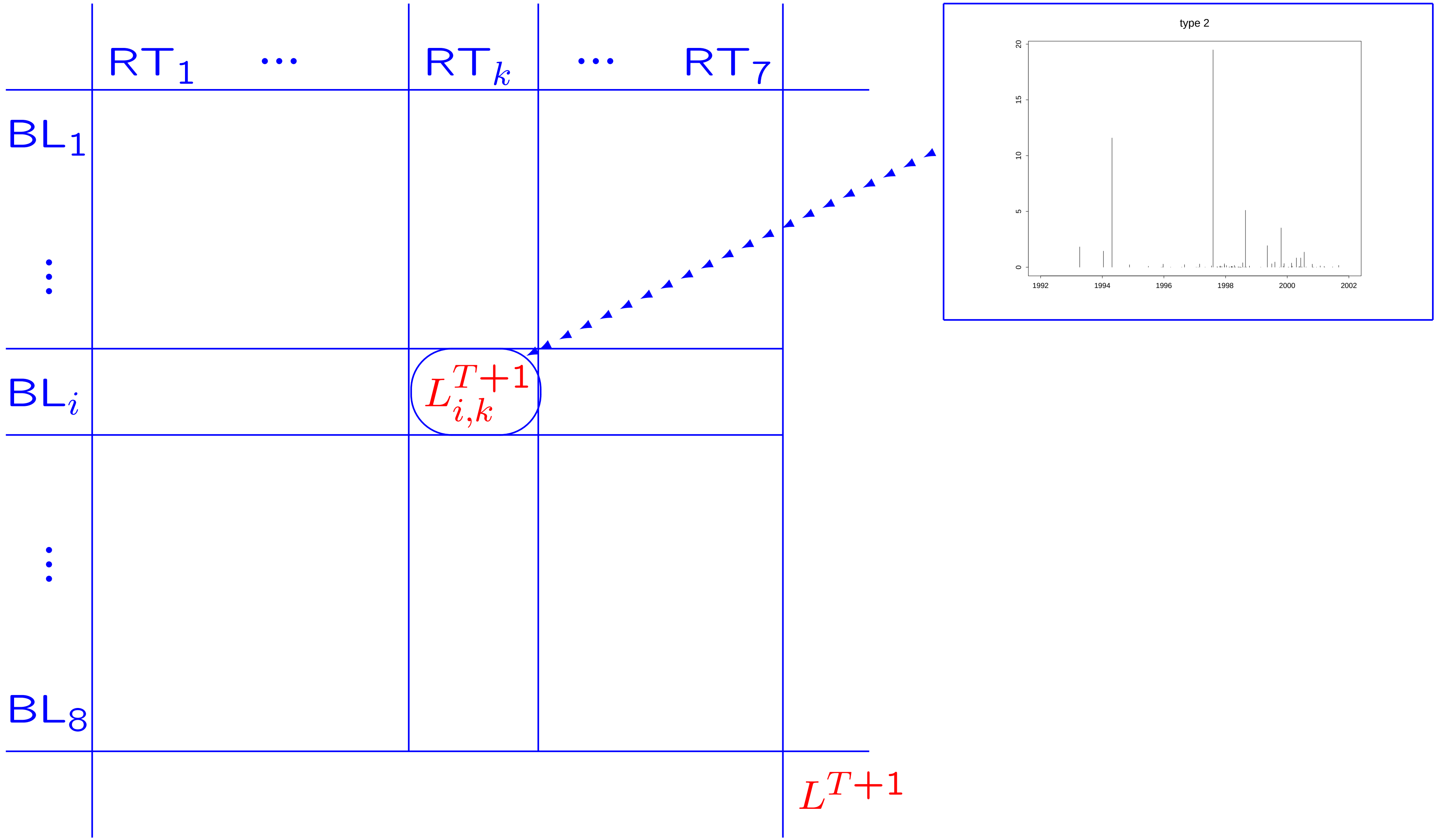
- Pillar 1 in LDA based on $\text{VaR}_{99.9\%}^{1\text{ year}}$, i.e. a 1 in 1000 year event!

- Various approaches are possible/allowed (Basel II spirit)

- Take care of: loss frequency, severity, lower truncation, and use internal/external/expert data

- insurance (up to 20%) is allowed

Loss Distribution Approach (LDA)



A complicated stochastic structure

$$L^{T+1} = \sum_{i=1}^{8} \sum_{k=1}^{7} L_{i,k}^{T+1}$$

$$L_{i,k}^{T+1} = \sum_{\ell=1}^{N_{i,k}^{T+1}} X_{i,k}^{\ell}$$

$X_{i,k}^{\ell}$: loss severities

$N_{i,k}^{T+1}$: loss frequencies

As an example, aggregate BL–wise, yielding risk capital estimates:

$$\mathrm{VaR}(1), \ldots, \mathrm{VaR}(8)$$

Calculate (as indicated in Basel **II**):

$$\mathrm{VaR}(+) = \mathrm{VaR}(1) + \cdots + \mathrm{VaR}(8)$$

Use diversification/corrrelation arguments to find the "real" capital charge

$$\mathrm{VaR} < \mathrm{VaR}(+)$$

Question: how reliable is the above procedure?

Here are some of the main issues:

- reliability of VaR(i) estimates (statistical issue)
- which arguments lead to VaR(+) reduction? (understanding correlation/diversification)
- are there practically relevant cases where

$$\text{VaR} > \text{VaR}(+) \quad ?$$

  (non–coherence of VaR)
- what about different risk measures

A related question: aggregating risk capital at the level of the bank

• MR:	$\mathrm{VaR}_{\alpha_1}^{d_1}$	$d_1 = 10$ days	$\alpha_1 = 99\%$
• CR:	$\mathrm{VaR}_{\alpha_2}^{d_2}$	$d_2 = 1$ year	$\alpha_2 = 99.9\%$
• OR:	$\mathrm{VaR}_{\alpha_3}^{d_3}$	$d_3 = 1$ year	$\alpha_3 = 99.9\%$
• EC:	$\mathrm{VaR}_{\alpha_4}^{d_4}$	$d_4 = 1$ year	$\alpha_4 = 99.97\%$

Calculate $\mathrm{VaR}_{\alpha}^{d}$ (Total) for given confidence level α and holding period d and link to EC

Problems to solve:

- scale $\text{VaR}_{\alpha_i}^{d_i}$ to a common $\text{VaR}_{\alpha}^{d}(i)$ for risk class $i = 1, 2, 3$

- calculate $\text{VaR}_{\alpha}^{d}(1)+\text{VaR}_{\alpha}^{d}(2)+\text{VaR}_{\alpha}^{d}(3)$

- discuss "diversification reduction"

Back to the concrete case of Operational Risk

- How to estimate a capital charge VaR(i) for BL i?
  - Using EVT ($\alpha = 99.9\%!$) as in Moscadelli, 2004?
  - Using the g–and–h distribution (r.v.)

    $$Y = A\frac{e^{gZ} - 1}{g}\, e^{\frac{1}{2}hZ^2} + B\,, \quad Z \sim N(0,1)$$

    as in Dutta and Perry, 2006?
  - Other methods?

- Key drivers towards choice (five “dimensions” in Dutta and Perry):
  - good fit to the available data
  - realistic wrt capital charge
  - well specified wrt data properties (e.g. losses are positive)
  - flexible wrt possible shape of loss densities
  - simple: “As simple as possible but not simpler”

- The VaR $< \sum_{i=1}^{8}$ VaR(i) issue is related to the coherence/non–coherence (in particular sub/superadditivity) of VaR:

$$\text{VaR}(X+Y) \leq \text{versus} \geq \text{VaR}(X) + \text{VaR}(Y)$$

$\geq$ may happen, even in practice

$=$ holds for comonotonic risks

$X = \Psi_1(W)\,,\quad Y = \Psi_2(W)\,,\quad \Psi_i$ increasing, $\quad W$ a rv

also yields maximal correlation between X and Y

- An important result: if $(X_1, \ldots, X_d)$ are multivariate normal then VaR is subadditive (coherent), even more generally this result holds for all multivariate elliptical vectors (e.g. t, logistic, hyperbolic, ...)

$$\mathbf{X} = \boldsymbol{\mu} + \sqrt{V}\mathbf{Z}\,, \quad \mathbf{Z} \sim N_d(\mathbf{0}, \boldsymbol{\Sigma}) \text{ independent of } V > 0$$

Why does this hold: look at the normal case

– $X \sim N\left(\mu, \sigma^2\right) \Rightarrow \text{VaR}_\alpha(X) = \mu + \sigma\Phi^{-1}(\alpha)$ (easy!)
VaR is driven by standard deviation σ and hence coherent (Cauchy–Schwartz Inequality)

– $(X_1, X_2) \sim N_2(\boldsymbol{\mu}, \boldsymbol{\sigma}, \rho)$, then $X_1 + X_2 \sim N\left(\mu_1 + \mu_2, \sigma_{1,2}^2\right)$
where $\sigma_{1,2}^2 = \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2$, hence

$$
\begin{aligned}
\text{VaR}_\alpha(X_1 + X_2) &= \mu_1 + \mu_2 + \sigma_{1,2}\Phi^{-1}(\alpha) \\
&= \mu_1 + \mu_2 + \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\,\Phi^{-1}(\alpha) \\
(\rho \leq +1) \quad &\leq \mu_1 + \mu_2 + \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2}\,\Phi^{-1}(\alpha) \\
&= \mu_1 + \mu_2 + (\sigma_1 + \sigma_2)\,\Phi^{-1}(\alpha) \\
&= \text{VaR}_\alpha(X_1) + \text{VaR}_\alpha(X_2)
\end{aligned}
$$

standard deviation drives subadditivity!

Under what circumstances is VaR superadditive ($>$)

– very skew distributions

– very heavy–tailed distributions

– special dependence but possible nice marginals

Are these examples relevant for practice?

- for MR perhaps not

- for CR slightly for very skew positions

- for OR highly relevant

  OpRisk data are:
  - heavy–tailed (power laws, Pareto)
  - skew
  - we have no clear view on interdependence

So "There is an issue!"

Here are the (counter)examples in short:

- Skewness
  100 iid loans: 2% yearly coupon, 100 face value, 1% yearly default, no recoverable, P&L X_i (use easy binomial calculation):

$$\text{VaR}_{95\%}\left(\sum_{i=1}^{100} X_i\right) > \text{VaR}_{95\%}(100X_1) = \sum_{i=1}^{100} \text{VaR}_{95\%}(X_i)$$

- Heavy–tailedness
  X_1, X_2 independent, $P(X_i > x) = x^{-\delta}L(x)$ with $\delta < 1$ (infinite mean!) and L slowly varying, then for α large enough

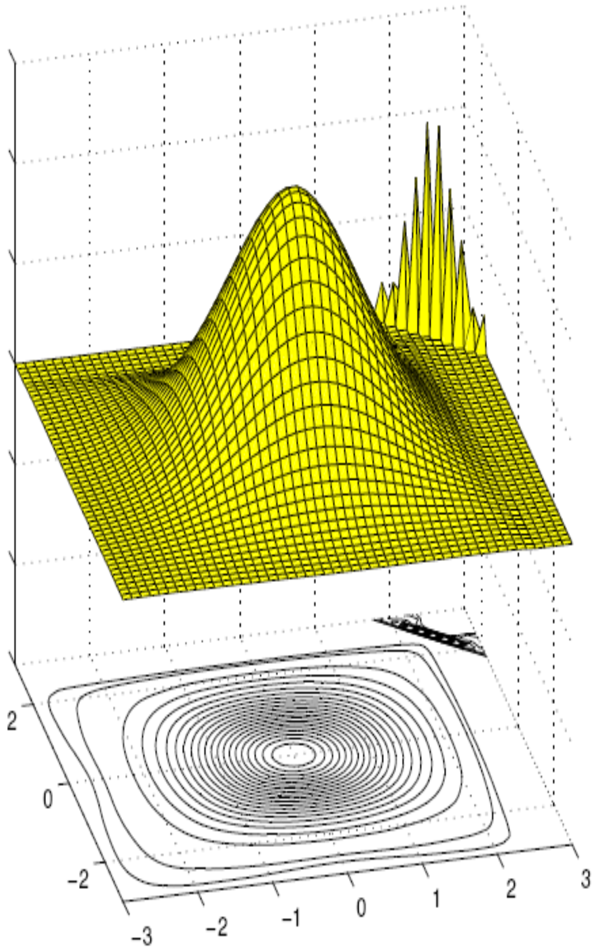
$$\text{VaR}_{\alpha}(X_1 + X_2) > \text{VaR}_{\alpha}(X_1) + \text{VaR}_{\alpha}(X_2)$$

- Special dependence

  given marginal risks $X_1 \sim F_1$, $X_2 \sim F_2$, then one can always construct a joint model F with marginals F_1, F_2 so that under F and for any $0 < \alpha < 1$:

$$\mathrm{VaR}_\alpha(X_1 + X_2) > \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$$

  use the notion of copula $F = C(F_1, F_2)$, in particular one can take $F_1 = F_2 = N(0,1)$ and construct the worst dependence structure with respect to VaR



“A risk manager’s worst nightmare”

Going beyond VaR, does that help?

– $\mathrm{VaR}_\alpha(X) = \inf\{\ell \in \mathbb{R} : F_X(\ell) \geq \alpha\} = F_X^{-1}(\alpha)$ (a quantile)

– Expected Shortfall

$$\begin{aligned} ES_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(X)\,du \qquad (1) \\ &= E\left(X \mid X \geq \mathrm{VaR}_\alpha(X)\right) \qquad \text{(careful!)} \end{aligned}$$

is always coherent ((1)), larger than VaR and yields information on the residual loss potential above VaR

– Median Shortfall

$\mathrm{Med}_\alpha(X) = \mathrm{median}\left(X \mid X \geq \mathrm{VaR}_\alpha(X)\right)$

may yield reasonable values but is in general not coherent

Conclusion (for the Operational Risk case)

- estimating $\text{VaR}_\alpha(1), \ldots, \text{VaR}_\alpha(d)$ for $\alpha = 0.999$ is a very daunting task: no easy solution!

- VaR(+) need not yield an upper limit!
  Hence discuss diversification effect with care

- No easy short–cut via scaling (90% $\rightarrow$ 99.9%), or constrained estimation (finite mean) or putting an upper limit on the losses

- Combining data (internal, external, expert opinion) may call for credibility techniques from actuarial science

- g–and–h seems promising, be careful with EVT

Back to the calculation of

$$\mathcal{R}\left(\Psi(\mathbf{X})\right) \tag{2}$$

for a risk/pricing measure $\mathcal{R}$, financial position Ψ and risk factors $\mathbf{X} = (X_1, \ldots, X_d)$

Suppose given marginal information $X_i \sim F_i$, $i = 1, \ldots, d$ and some idea of dependence, then (2) cannot be calculated explicitly, hence approximations/bounds are called for

$$\mathcal{R}_L \leq \mathcal{R}(\Psi(\mathbf{X})) \leq \mathcal{R}_U$$

Tasks: Calculate $\mathcal{R}_L$, $\mathcal{R}_U$

Prove optimality

Some concrete examples:

- homogeneous case ($F_1 = \cdots = F_8$) (easier)

$$P\,(X_i > x) = (1+x)^{-1.5}\,, \quad x \geq 0\,, \quad i = 1,\ldots,8$$

  - comonotonic case

$$\mathrm{VaR}_{0.999}\left(\sum_{i=1}^{8} X_i\right) = \sum_{i=1}^{8} \mathrm{VaR}_{0.999}\,(X_i) = 0.79$$

  - $\mathcal{R}_U = 1.93$

- heterogeneous case ($F_1 \neq \cdots \neq F_8$) (difficult)
  $P\,(X_i > x)$ as given by Moscadelli (2004)
  - Comonotonic case $\sum_{i=1}^{8} \mathrm{VaR}_{0.999}\,(X_i) = 4.8347 \times 10^5$
  - $\mathcal{R}_U = 2.3807 \times 10^6$ (factor 5!)

Final conclusions

- Quantitative tools exist for understanding aggregation issues
- In special cases, calculation is easy
- In general difficult
- Diversification effects have to be handled with care
- Definitely more research is needed
- Basel II still has some issues to solve (2007+)

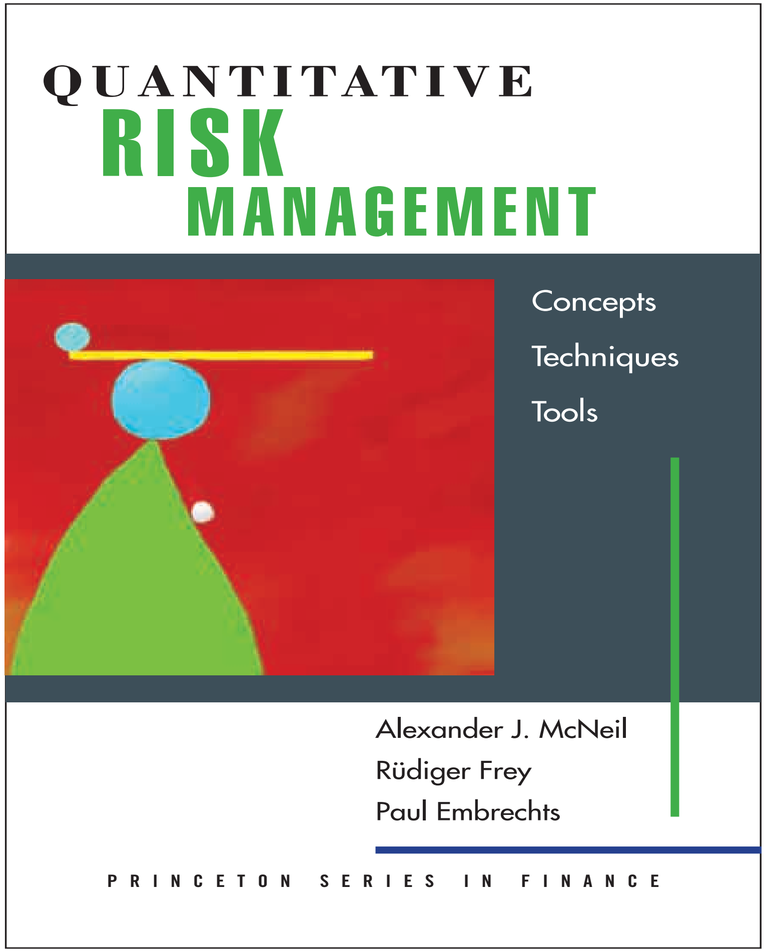
- VaR–aggregation leads to problems

- Scaling is only understood in very specific models

- And as always

MORE WORK IS NEEDED

Further reading



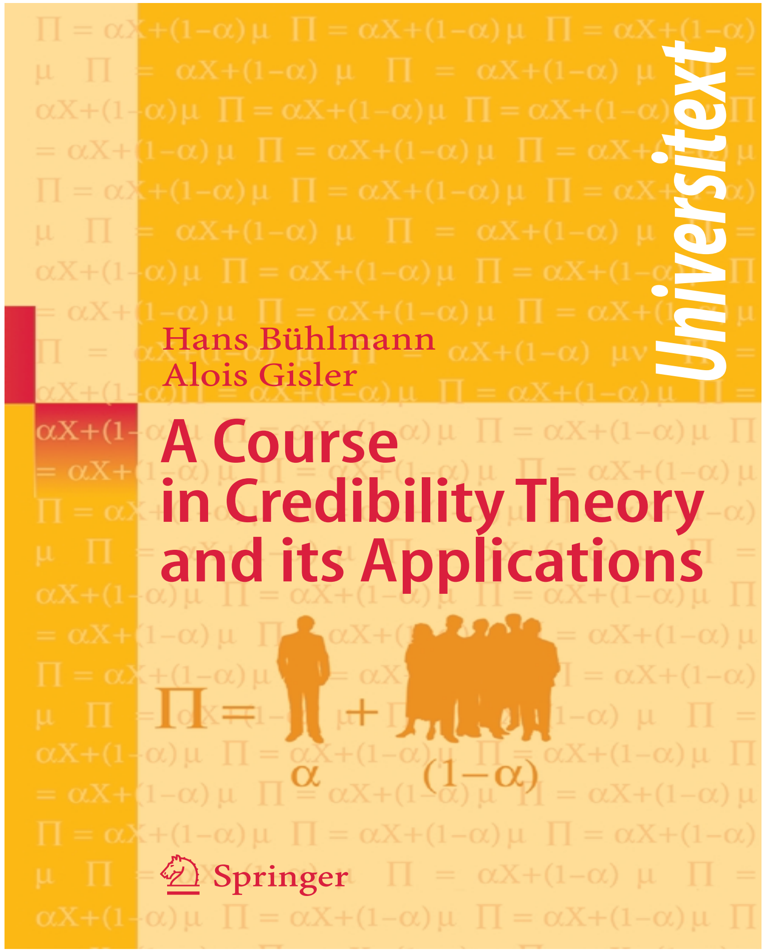


Table of Content (538 pages)

1. Risk in Perspective

2. Basic Concepts in Risk Management

3. Multivariate Methods

4. Financial Time Series

5. Copulas and Dependence

6. Aggregate Risk

7. Extreme Value Theory

8. Credit Risk Management

9. Dynamic Credit Risk Models

10. Operational Risk and Insurance Analytics

Visit the Book Zone:

http://www.ma.hw.ac.uk/~mcneil/book/index.html