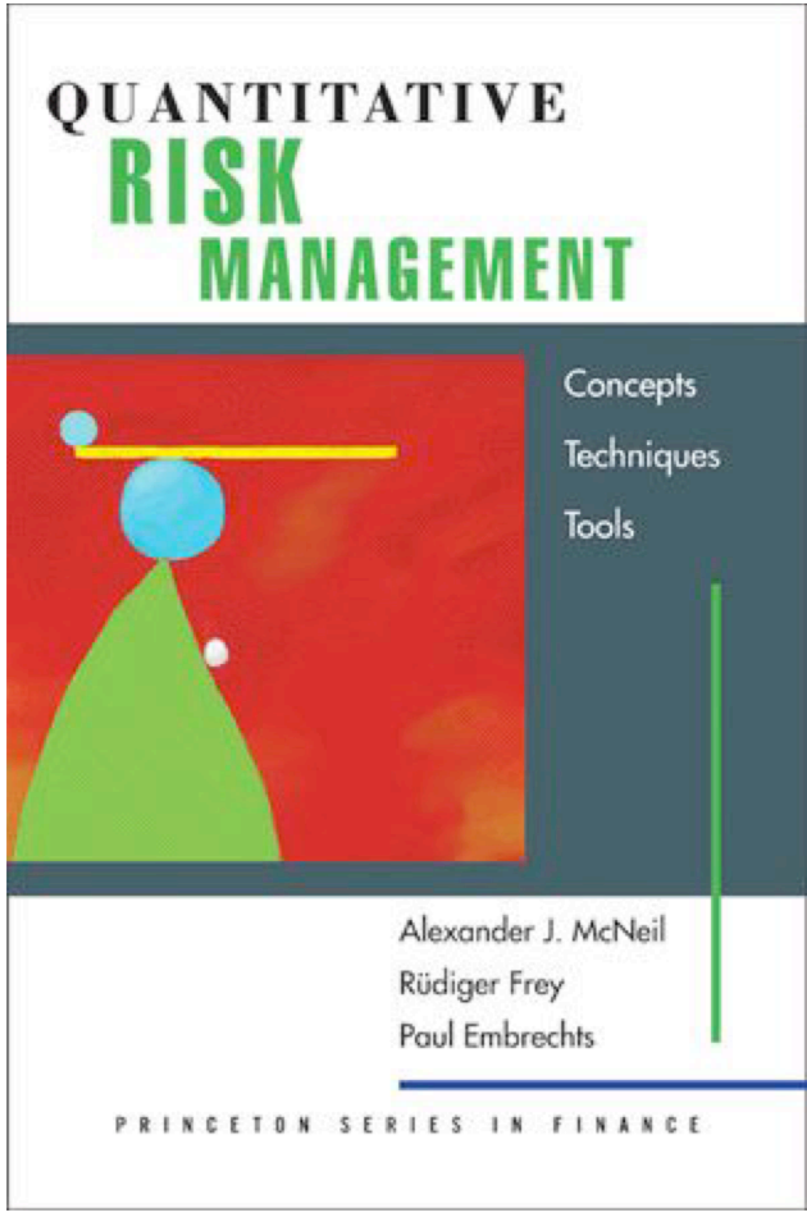
Quantitative Methods for Risk Management

Paul Embrechts and Johanna Nešlehová
ETH Zurich
www.math.ethz.ch/∼embrecht

BaFin 2006


QUANTITATIVE
RISK
MANAGEMENT
Concepts
Techniques
Tools
Alexander J. McNeil
Rüdiger Frey
Paul Embrechts
PRINCETON SERIES IN FINANCE

L. Operational Risk and Insurance Analytics

1. A New Risk Class
2. Insurance Analytics Toolkit
3. The Capital Charge Problem
4. Marginal VaR Estimation
5. Global VaR Estimation

L1. A New Risk Class

The New Accord (Basel II)

- **1988**: Basel Accord (Basel I): minimal capital requirements against credit risk, one standardised approach, Cooke ratio
- **1996**: Amendment to Basel I: market risk, internal models, netting
- **1999**: Several Consultative Papers on the New Accord (Basel II)
- **to date**: CP3: Third Consultative Paper on the New Basel Capital Accord (`www.bis.org/bcbs/`)
- **2007+**: full implementation of Basel II

Basel II: What is new?

- **Rationale** for the New Accord: More flexibility and risk sensitivity

- **Structure** of the New Accord: Three-pillar framework:

❶ Pillar 1: minimal capital requirements (risk measurement)

❷ Pillar 2: supervisory review of capital adequacy

❸ Pillar 3: public disclosure

Basel II: What is new? (cont'd)

- Two options for the measurement of credit risk:
  - Standard approach
  - Internal rating based approach (IRB)

- Pillar 1 sets out the minimum capital requirements (Cooke Ratio, McDonough Ratio):

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

- MRC (minimum regulatory capital) $\stackrel{\text{def}}{=}$ 8% of risk-weighted assets

- Explicit treatment of operational risk

Operational Risk

Definition:

The risk of losses resulting from inadequate or failed internal processes, people and systems, or external events.

Remark:

This definition includes legal risk, but excludes strategic and reputational risk.

Note:

Solvency 2

Operational Risk (cont'd)

- **Notation**: C_{OP}: capital charge for operational risk
- **Target**: $C_{\mathsf{OP}} \approx 12\%$ of MRC (down from initial 20%)
- **Estimated total losses** in the US (2001): $50b
- **Some examples**
  - 1977: Credit Suisse Chiasso-affair
  - 1995: Nick Leeson/Barings Bank, £1.3b
  - 2001: September 11
  - 2001: Enron (largest US bankruptcy so far)
  - 2002: Allied Irish, £450m

Risk Measurement Methods for Operational Risk

Pillar 1 regulatory minimal capital requirements for operational risk:

Three distinct approaches:

1. Basic Indicator Approach

2. Standardised Approach

3. Advanced Measurement Approach (AMA)

Basic Indicator Approach

- Capital charge:

$$\boxed{C_{\mathsf{OP}}^{\mathsf{BIA}} = \alpha \times GI}$$

- $C_{\mathsf{OP}}^{\mathsf{BIA}}$: capital charge under the Basic Indicator Approach

- GI: average annual gross income over the previous three years

- $\alpha = 15\%$ (set by the Committee based on CISs)

Standardised Approach

- Similar to the BIA, but on the level of each business line:

$$\boxed{C_{\text{OP}}^{\text{SA}} = \sum_{i=1}^{8} \beta_i \times GI_i}$$

$\beta_i \in [12\%, 18\%]$, $i = 1, 2, \ldots, 8$ and 3-year averaging

- 8 business lines:

Corporate finance (18%)	Payment & Settlement (18%)
Trading & sales (18%)	Agency Services (15%)
Retail banking (12%)	Asset management (12%)
Commercial banking(15%)	Retail brokerage (12%)

Advanced Measurement Approach (AMA)

- Allows banks to use their internally generated risk estimates
- Preconditions: Bank must meet qualitative and quantitative standards before being allowed to use the AMA
- Risk mitigation via insurance possible ($\leq 20\%$ of $C_{\mathsf{OP}}^{\mathsf{SA}}$)
- Incorporation of risk diversification benefits allowed
- "Given the continuing evolution of analytical approaches for operational risk, the Committee is not specifying the approach or distributional assumptions used to generate the operational risk measures for regulatory capital purposes."
- Example: **Loss distribution approach**

Internal Measurement Approach

- Capital charge (similar to Basel II model for Credit Risk):

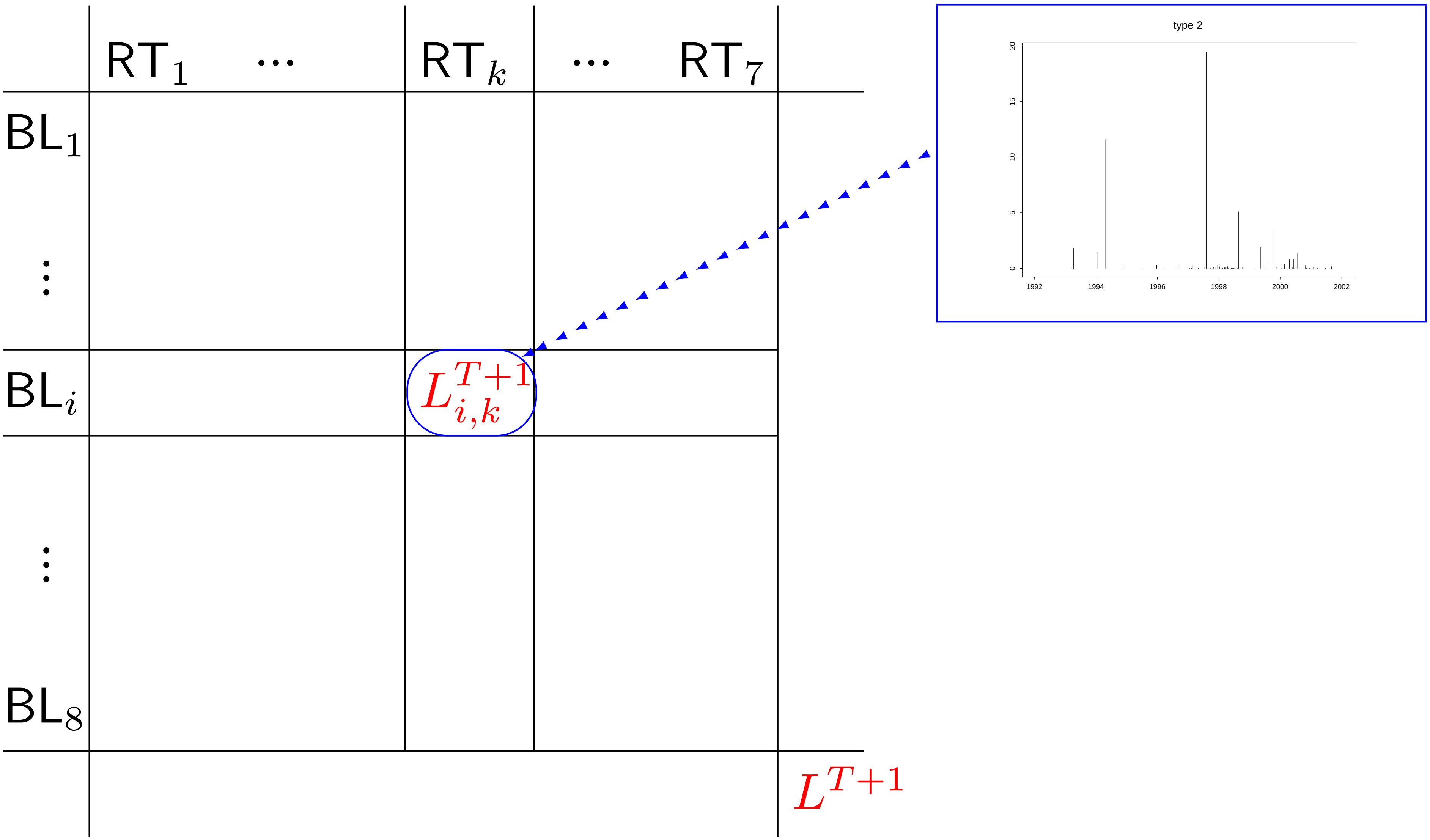
$$\boxed{C_{\text{OP}}^{\text{IMA}} = \sum_{i=1}^{8} \sum_{k=1}^{7} \gamma_{ik}\, e_{ik}} \qquad \text{(first attempt)}$$

e_{ik}: expected loss for business line i, risk type k

γ_{ik}: scaling factor

- 7 loss types:
  Internal fraud
  External fraud
  Employment practices and workplace safety
  Clients, products & business practices
  Damage to physical assets
  Business disruption and system failures
  Execution, delivery & process management

Loss Distribution Approach (LDA)



LDA: continued

- For each business line/loss type cell (i,k) one models

$L_{i,k}^{T+1}$: OP risk loss for business line i, loss type k over the future (one year, say) period $[T, T+1]$

$$L_{i,k}^{T+1} = \sum_{\ell=1}^{N_{i,k}^{T+1}} X_{i,k}^{\ell} \quad \text{(next period's loss for cell } (i,k))$$

Note that $X_{i,k}^{\ell}$ is truncated from below

LDA: continued

Remark: Look at the structure of the loss random variable L^{T+1}

$$
\begin{aligned}
L^{T+1} &= \sum_{i=1}^{8} \sum_{k=1}^{7} L_{i,k}^{T+1} \quad \text{(next period's total loss)} \\
&= \sum_{i=1}^{8} \sum_{k=1}^{7} \sum_{\ell=1}^{N_{i,k}^{T+1}} X_{i,k}^{\ell} \\
&= \sum_{i=1}^{8} L_i^{T+1} \quad \text{(often used decomposition)}
\end{aligned}
$$

- Check again the overall complexity of the (BL, RT) matrix

L2. Insurance Analytics: an Essential Toolkit

Total Loss Amount

Denote by $N(t)$ the (random) number of losses over a fixed period $[0, t]$ and write $X_1, X_2, \ldots$ for the individual losses. The aggregate loss is

$$S_{N(t)} = \sum_{k=1}^{N(t)} X_k$$

Remarks:

- $F_{S_{N(t)}}(x) = P(S_{N(t)} \leq x)$ is called the total loss df. If t is fixed, we write S_N and F_{S_N} instead

- The random variable $S_{N(t)}$ is also referred to as random sum

Compound Sums

Assume:

1. (X_k) are iid with common df G, $G(0) = 0$

2. N and (X_k) are independent

S_N is then referred to as a compound sum. The pdf of N is denoted by $p_N(k) = P(N = k)$, $k = 0, 1, \ldots$ and N is called a compounding rv.

Proposition 1:

Let S_N be a compound sum and the above assumptions hold. Then

$$F_{S_N}(x) = \sum_{k=0}^{\infty} p_N(k) G^{*k}(x), \quad x \geq 0$$

Proposition 2:

Let S_N be a compound sum and the above assumptions hold. Then the Laplace-Stieltjes transform of S_N satisfies

$$\hat{F}_{S_N}(s) = \int_0^\infty e^{-sx} dF_{S_N}(x) = \sum_{k=0}^\infty p_N(k)\hat{G}^k(s) = M_N(\hat{G}(s)), \quad s \geq 0$$

where M_N denotes the moment-generating function of N.

Proposition 3:

Let S_N be a compound sum and the above assumptions hold. If $E(N^2) < \infty$ and $E(X_1^2) < \infty$, we have that

$$\begin{aligned} E(S_N) &= E(N)E(X_1) \\ \mathrm{var}(S_N) &= \mathrm{var}(N)(E(X_1))^2 + E(N)\mathrm{var}(X_1) \end{aligned}$$

Compound Poisson Distribution

Example 1:

Consider $N \sim \text{Poi}(\lambda)$. Then S_N is referred to as a compound Poisson rv.

- The moment-generating function of N satisfies $M_N(s) = \exp(-\lambda(1-s))$ and hence

$$\hat{F}_{S_N}(s) = \exp(-\lambda(1-\hat{G}(s)))$$

- Notation: $S_N \sim \text{CPoi}(\lambda, G)$

- If $E(X_1^2) < \infty$, the moments of S_N are by Proposition 3

$$E(S_N) = \lambda E(X_1) \quad \text{and} \quad \text{var}(S_N) = \lambda E(X_1^2)$$

Aggregation of Compound Poisson rvs

Suppose that the compound sums $S_{N_i} \sim \text{CPoi}(\lambda_i, G_i)$, $i = 1, \ldots, d$ and that these rvs are independent. Then

$$S_N := \sum_{i=1}^{d} S_{N_i} = \sum_{i=1}^{d} \sum_{k=1}^{N_i} X_{i,k}$$

is again a compound Poisson rv, $S_N \sim \text{CPoi}(\lambda, G)$ where

$$\lambda = \sum_{i=1}^{d} \lambda_i \quad \text{and} \quad G = \sum_{i=1}^{d} \frac{\lambda_i}{\lambda} G_i$$

G is hence a mixture distribution. A simulation from G can be done in two steps: first draw i, $i \in \{1, \ldots, d\}$ with probability λ_i/λ and then draw a loss with df G_i.

Binomial Loss Model

Example 2:

Suppose $N \sim \text{Bin}(n, p)$. S_N is then called the (individual risk) binomial model.

Consider a time interval $[0, 1]$ and let N_n denote the total number of losses in $[0, 1]$ for a fixed n. Suppose further that we have a number of potential loss generators that can produce, with probability p_n, a loss in each small subinterval $((k-1)/n, k/n]$, $k = 1, \ldots, n$. Moreover, the occurrence of a loss in any particular subinterval is not influenced by the occurrence of losses in other intervals and $np_n \to \lambda$ for a $\lambda > 0$ as $n \to \infty$.

For a fixed severity distribution, S_{N_n} is then a binomial model with $N_n \sim \text{Bin}(n, p_n)$ and converges in law to a compound Poisson rv with parameter λ as $n \to \infty$.

Over-dispersion

For compound Poisson rvs with $N \sim \text{Poi}(\lambda)$ we have that

$$E(N) = \text{var}(N) = \lambda$$

- Count data however often exhibit over-dispersion meaning that they indicate $E(N) < \text{var}(N)$
- This can be achieved by mixing, i.e. by randomizing the parameter λ

Randomization

Other examples of mixing include:

- from Black-Scholes to stochastic volatility models (randomize σ)
- from $\mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ to elliptical distributions (randomize Σ); i.e. the multivariate t distribution. Randomization of both $\boldsymbol{\mu}$ and Σ leads to generalized hyperbolic distributions
- mixing models for credit risk (randomizing the default probability)
- credibility theory in insurance (randomizing the underlying risk parameter)
- Bayesian inference ...

Poisson Mixtures

Definition: Let Λ be a positive rv with distribution function F_Λ. A rv N given by

$$P(N=k)=\int_0^\infty P(N=k|\Lambda=\lambda)dF_\Lambda(\lambda)=\int_0^\infty e^{-\lambda}\frac{\lambda^k}{k!}dF_\Lambda(\lambda)$$

is called a mixed Poisson rv with structure or mixing distribution F_Λ. A compound sum with a mixed Poisson rv as the compounding rv is called a compound mixed Poisson rv.

Lemma: Suppose that N is mixed Poisson with structure df F_Λ. Then $E(N)=E(\Lambda)$ and $\mathrm{var}(N)=E(\Lambda)+\mathrm{var}(\Lambda)$, i.e. for Λ non-degenerate, N is over-dispersed.

Negative Binomial Distribution

Example 3: For $\Lambda \sim \mathsf{Ga}(\alpha, \beta)$, the mixed Poisson rv is negative binomial, $N \sim \mathsf{NB}(\alpha, \beta/(\beta+1))$:

$$P(N = k) = \left(\frac{\beta}{\beta+1}\right)^{\alpha} \left(\frac{1}{\beta+1}\right)^{k} \frac{\Gamma(\alpha+k)}{(\beta+1)^{\alpha+k}}$$

Further,

$$E(N) = \frac{\alpha}{\beta} = E(\Lambda)$$

$$\mathrm{var}(N) = \frac{\alpha(\beta+1)}{\beta^2} = \frac{\alpha}{\beta} + \frac{\alpha}{\beta^2} = E(\Lambda) + \mathrm{var}(\Lambda)$$

Compounding leads to compound mixed Poisson rvs.

- Many more interesting models exist.

Approximations

- The distribution of S_N is generally intractable

Normal approximation for CPoi(λ, G):

$$F_{S_N}(x) \approx \Phi\left(\frac{x - E(N)E(X_1)}{\sqrt{\mathrm{var}(N)(E(X_1))^2 + E(N)\mathrm{var}(X_1)}}\right)$$

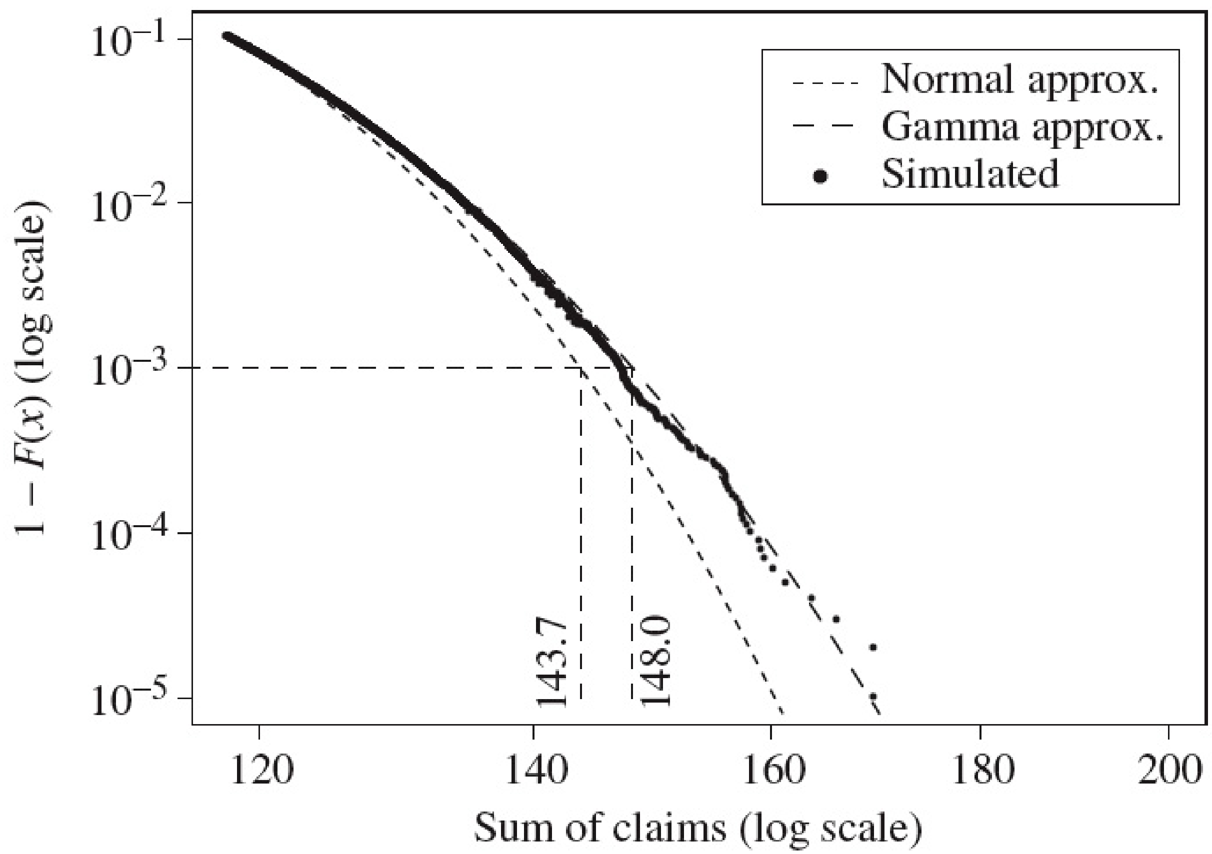
However, the skewness of S_N is positive:

$$\frac{E(S_N - E(S_N))^3}{(\mathrm{var}(S_N))^{3/2}} = \frac{E(X_1^3)}{\sqrt{\lambda}} > 0$$

Translated-gamma approximation for CPoi(λ, G):

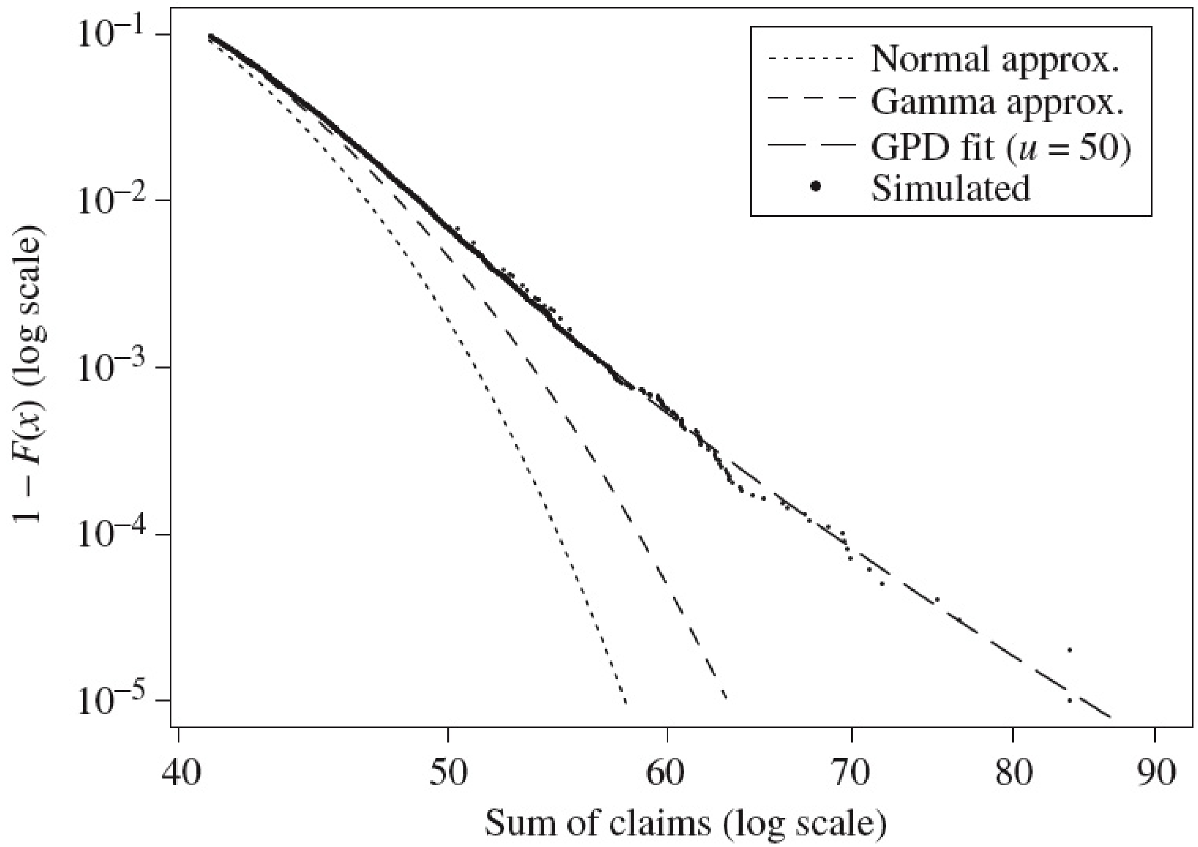
Approximate S_N by $k + Y$ where k is a translation parameter and $Y \sim$ Ga(α, β). The parameters k, α, β are found by matching the mean, variance and skewness.

Example: Light-tailed Severities



Simulated CPoi$(100,$ Exp$(1))$ data together with normal and translated gamma approximations. The 99.9% quantile estimates are also given.

Example: Heavy-tailed Severities



Simulated CPoi$(100, \text{Pa}(4, 1))$ data together with normal and translated gamma approximations. GPD approximation based on the POT method is also performed.

Panjer Class

- Recursive method for approximating S_N in case the severity distribution G is discrete and N satisfies a specific condition.

Panjer Class

The probability mass distribution of N belongs to the Panjer(a, b) class for some $a, b \in \mathbb{R}$ if $p_N(k) = (a + (b/k))p_N(k-1)$ for $k \geq 1$

The only nondegenerate examples of distributions belonging to a Panjer(a, b) class are

- binomial B(n, p) with $a = -p/(1-p)$ and $b = (n+1)p/(1-p)$
- Poisson Poi(λ) with $a = 0$ and $b = \lambda$
- Negative binomial NB(α, p) with $a = 1 - p$ and $b = (\alpha - 1)(1 - p)$

Panjer Recursion

For a discrete severity rv X_1 we denote

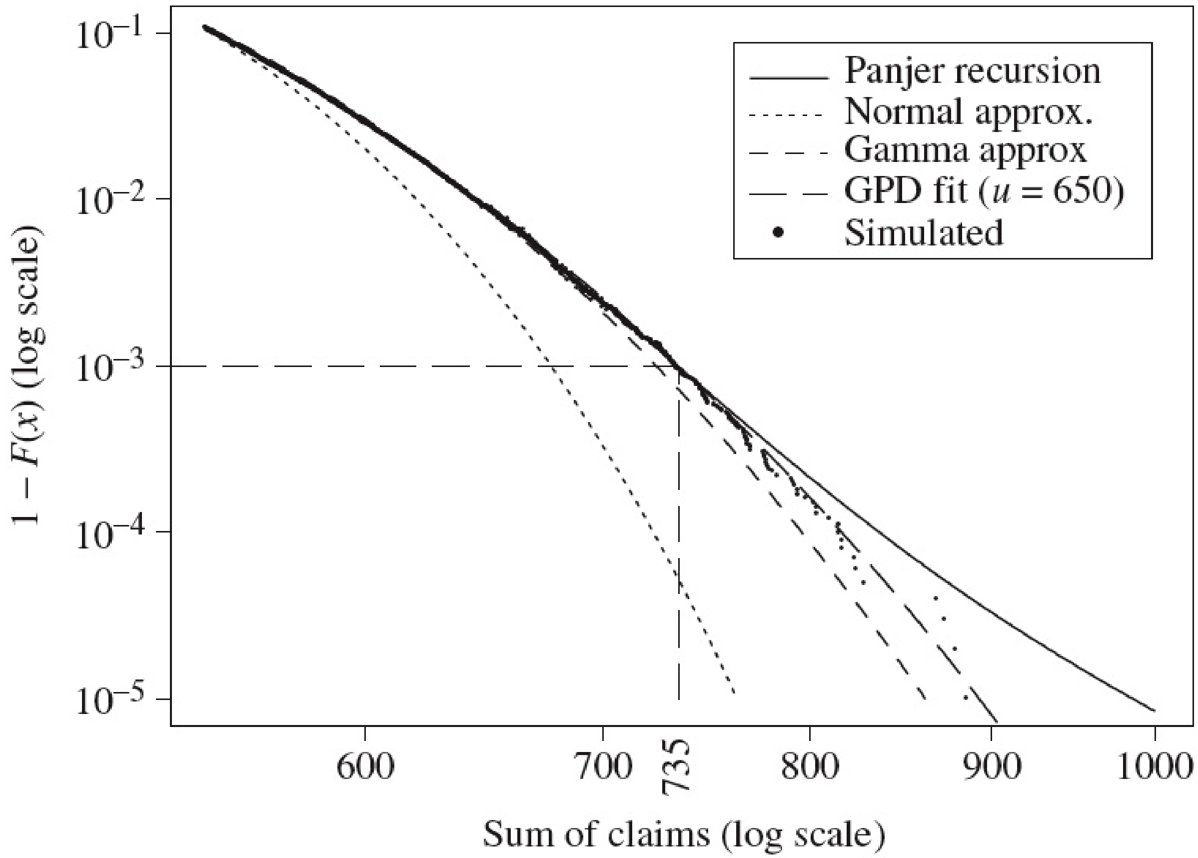
$$g_i := P(X_1 = i) \quad \text{and} \quad s_i := P(S_N = i)$$

Theorem: Suppose N satisfies the Panjer(a, b) class condition and $g_0 = 0$. Then $s_0 = p_N(0) = 0$ and, for $k \geq 1$,

$$s_k = \sum_{i=1}^{k} \left(a + \frac{bi}{k} \right) g_i s_{k-i}.$$

- For continuous severity distributions, discretization necessary
- Correction for $g_0 > 0$ possible
- Estimation of s_k far in the tail is more tricky

Example



Simulated CPoi$(100, \text{LN}(1,1))$ data together with the Panjer recursion approximation. Normal, translated gamma and GPD approximations are also performed.

Further Topics

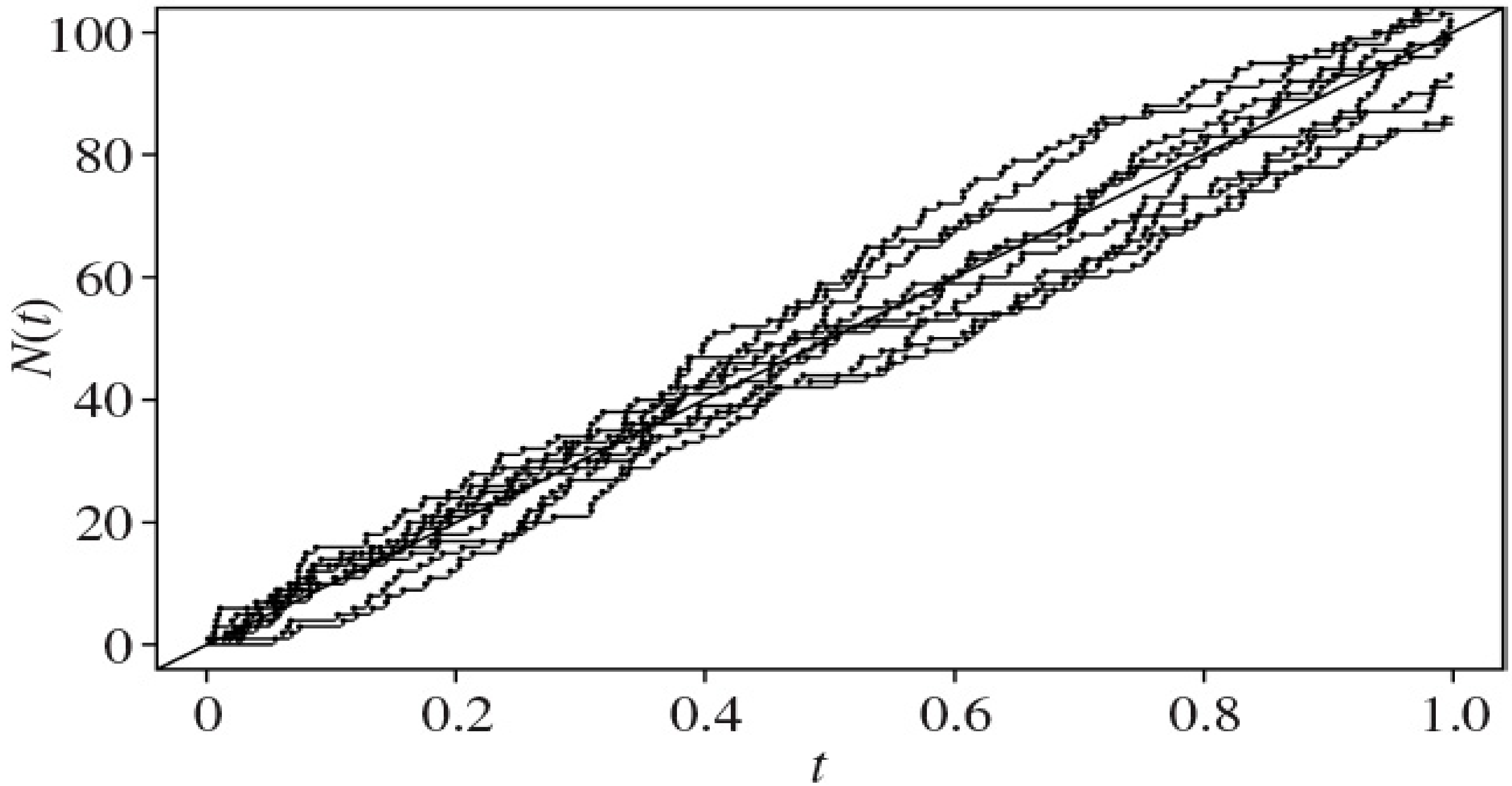
- S_N can be looked upon as a process in time, $S_{N(t)}$. Instead of N we then have the process $\{N(t), t \geq 0\}$ counting the number of events in $[0, t]$. Interesting examples for $N(t)$ are
  - homogeneous Poisson process
  - non-homogeneous Poisson process
  - Cox or doubly stochastic Poisson process

- Of further interest is the surplus process $C_t = u + ct - S_{N(t)}$ and the corresponding ruin probability

$$\psi(u, T) = P^u\{\inf_{0 \leq t \leq T} C_t \leq 0\}$$

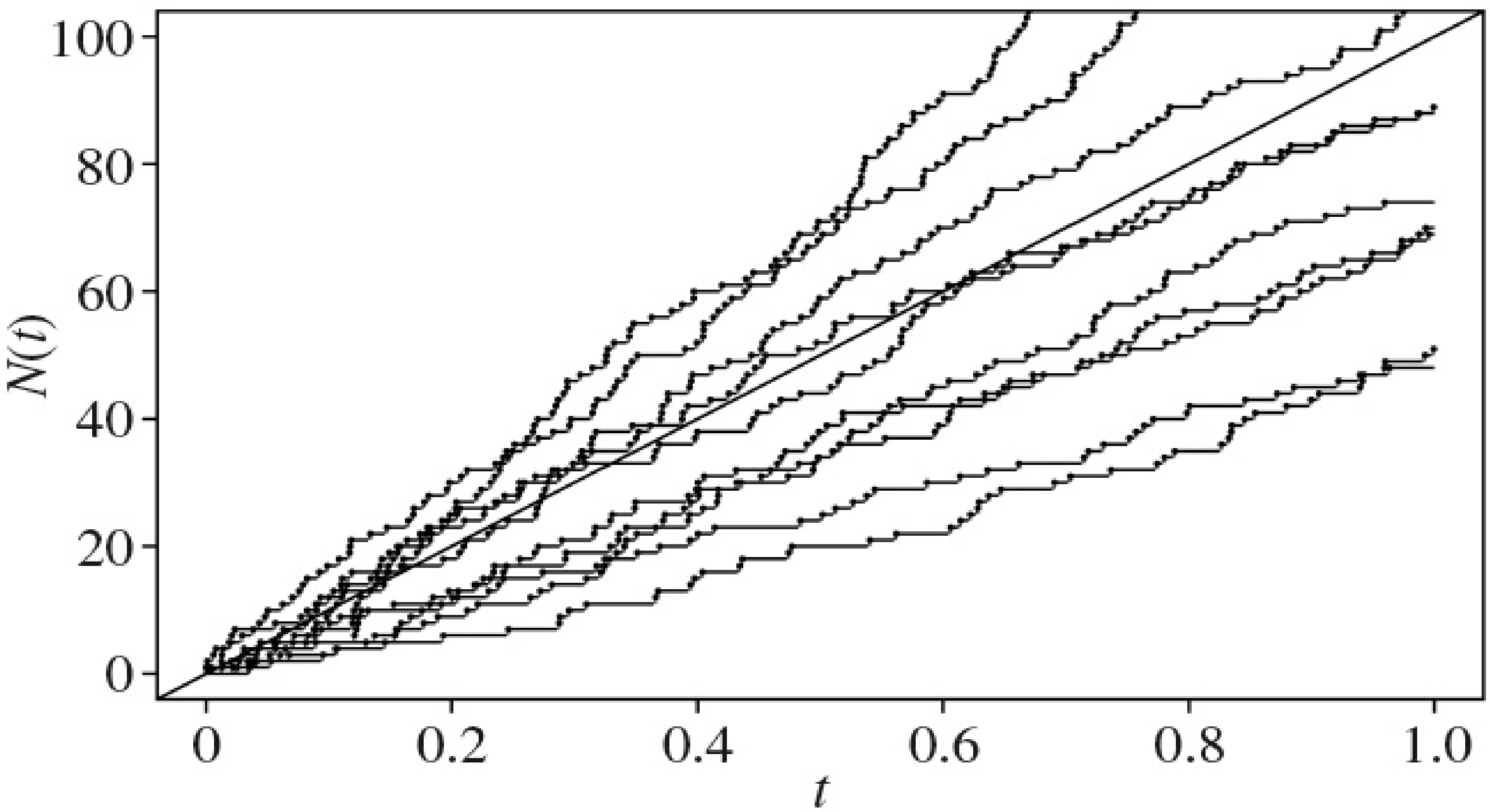
- Rare event simulation

Homogeneous Poisson Process



Ten realizations of a homogeneous Poisson process with $\lambda = 100$.

Mixed Poisson Process



Ten realizations of a mixed Poisson process with $\Lambda \sim \text{Ga}(100, 1)$.

Ruin Probability

Recall ruin results for

$$\psi(u) := \psi(u,\infty) = P^u\{\inf_{0\le t\le\infty} C_t < 0\}$$

- **Cramér-Lundberg**: "small claims"

$$\psi(u) < e^{-Ru}, \quad \forall u > 0$$
$$\psi(u) \sim \epsilon_1 e^{-Ru}, \quad u \to \infty$$

- **Embrechts-Veraverbeke**: "large claims" with df G

$$\psi(u) \sim \epsilon_2 \int_u^\infty \overline{G}(t)dt, \quad u \to \infty$$

Subexponential Distributions

Definition:

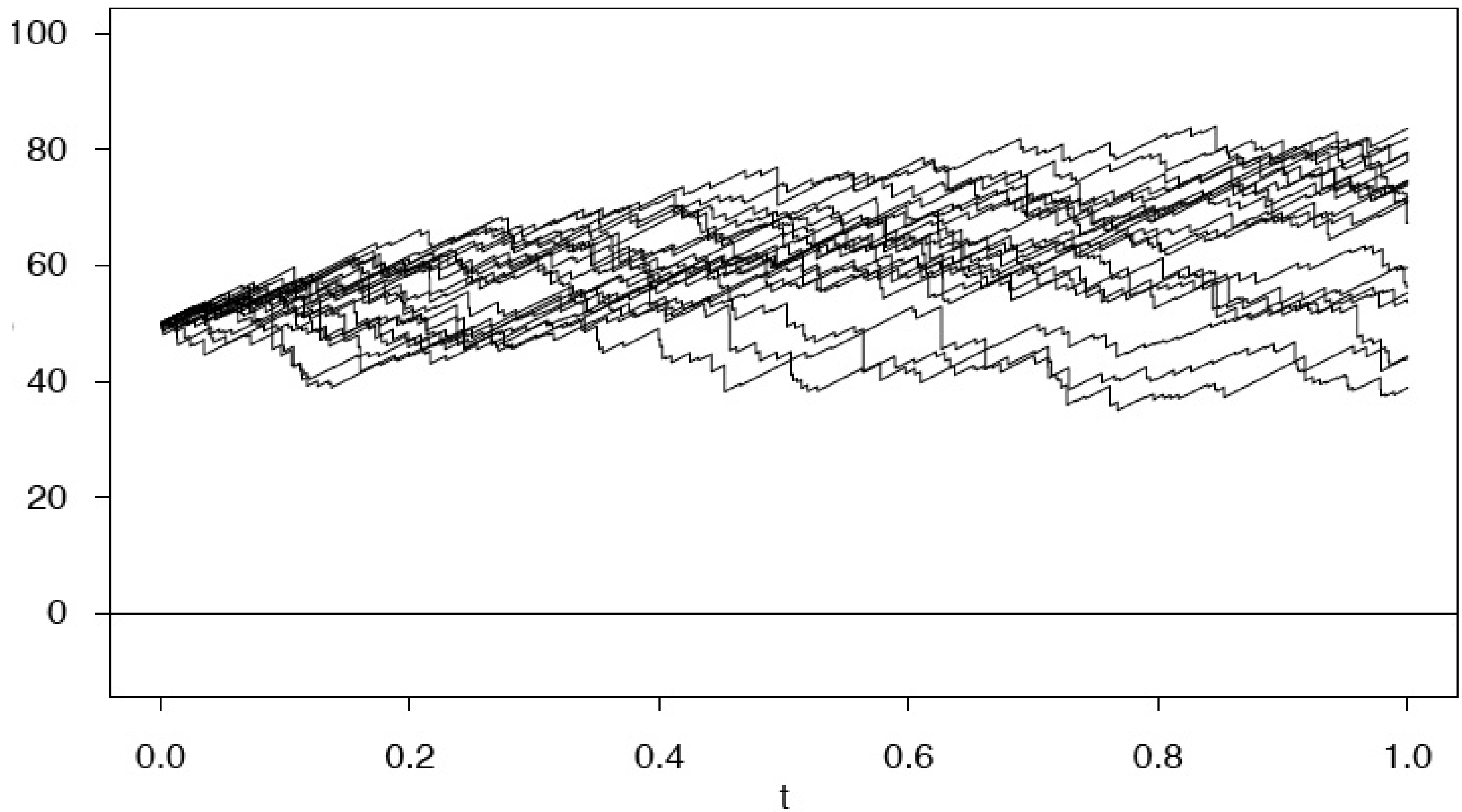
For $X_1, \ldots, X_n$ positive iid random variables with common distribution function F_X, denote $S_n = \sum_{k=1}^n X_k$ and $M_n = \max(X_1, \ldots, X_n)$. The distribution function F_X is called subexponential (denoted by $F_X \in \mathcal{S}$) for some (and then for all) $n \geq 2$ if

$$\lim_{x \to \infty} \frac{P(S_n > x)}{P(M_n > x)} = 1$$

Examples:

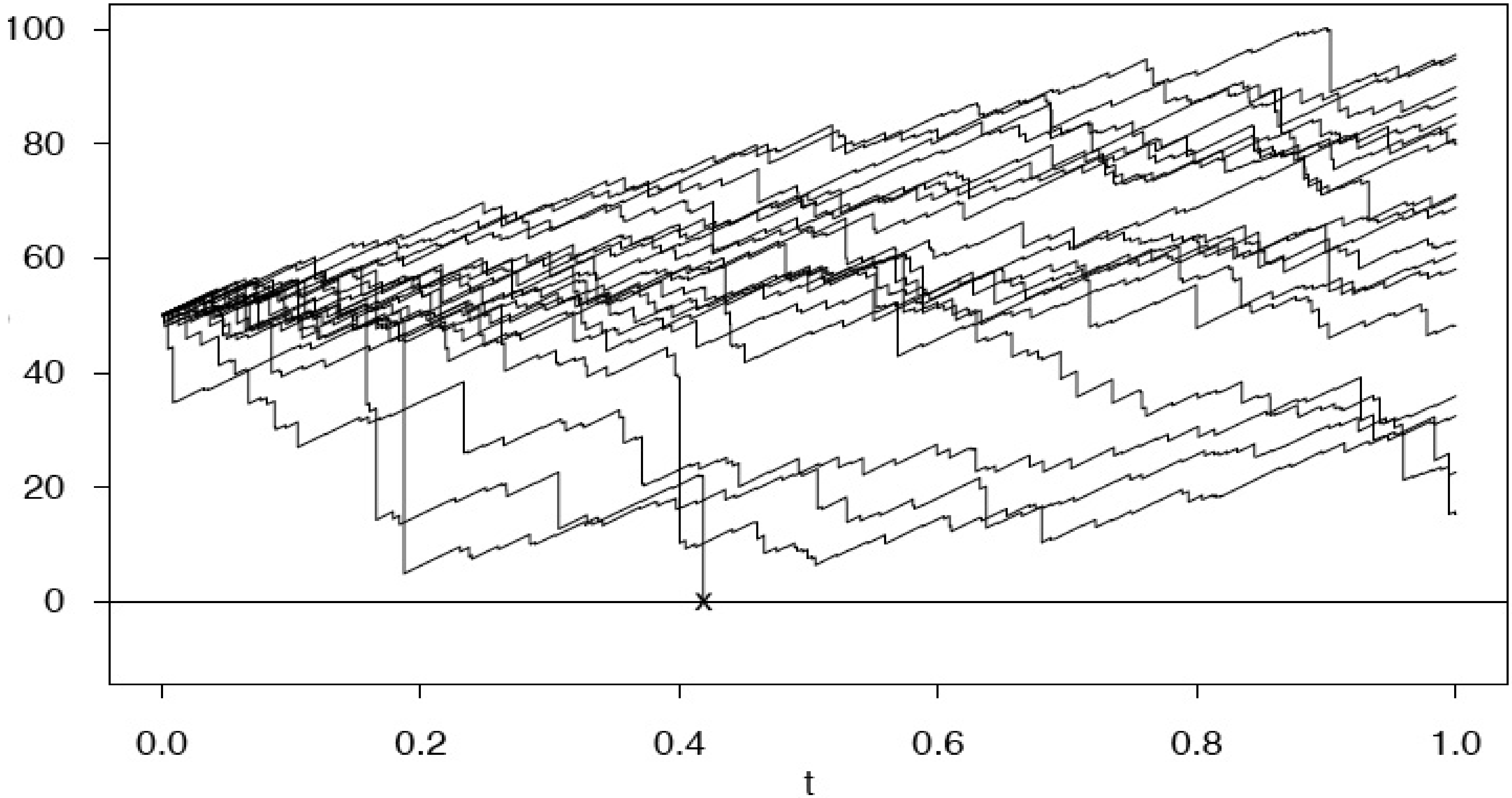
- Pareto, Generalized Pareto, Lognormal, Loggamma, ...

Ruin Process with Exponential Claims



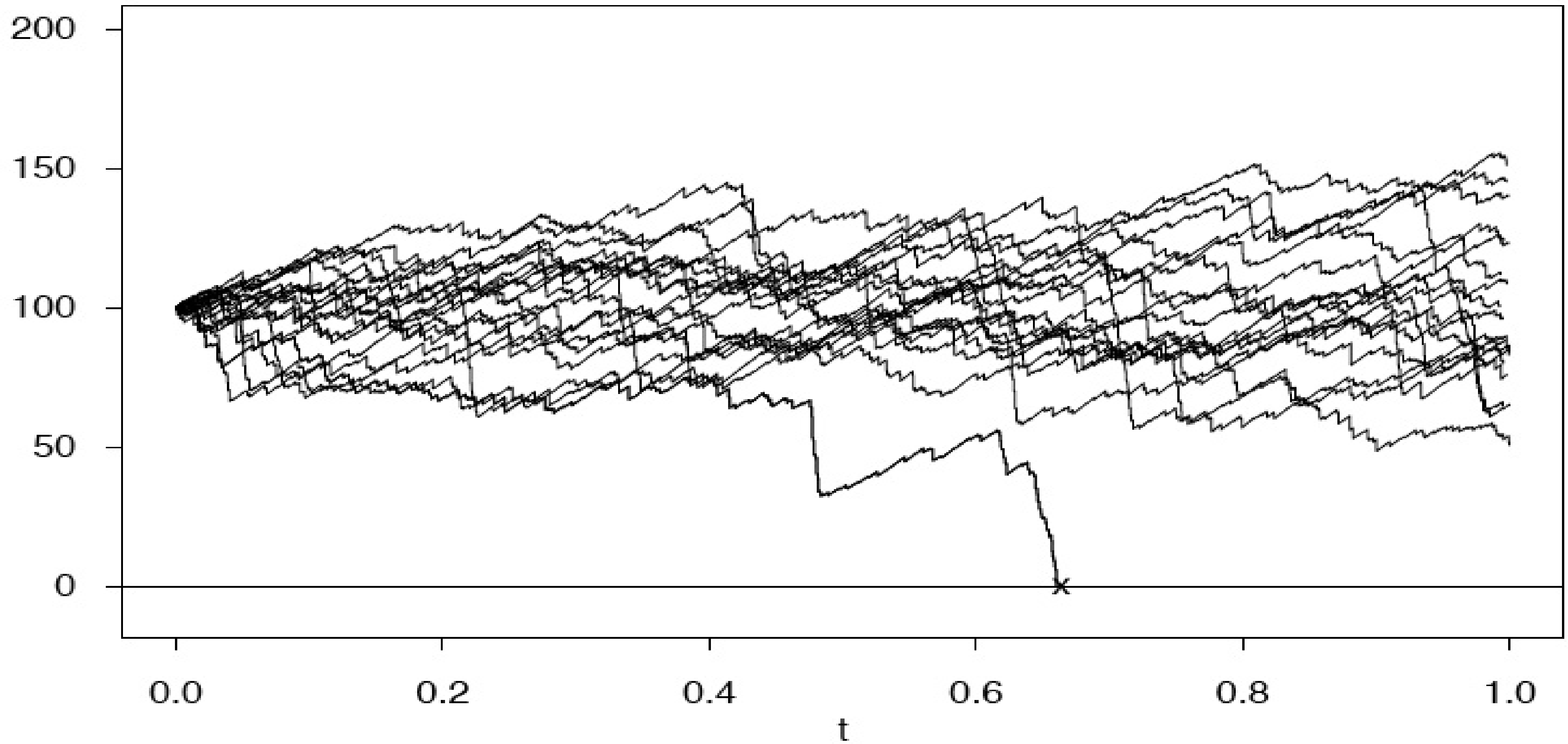
20 simulations from the ruin process C_t, $0 \le t \le 1$, with $(N(t)) \sim \text{HPois}(100t)$ and $X_1 \sim \text{Exp}(1)$.

Ruin Process with Pareto Claims



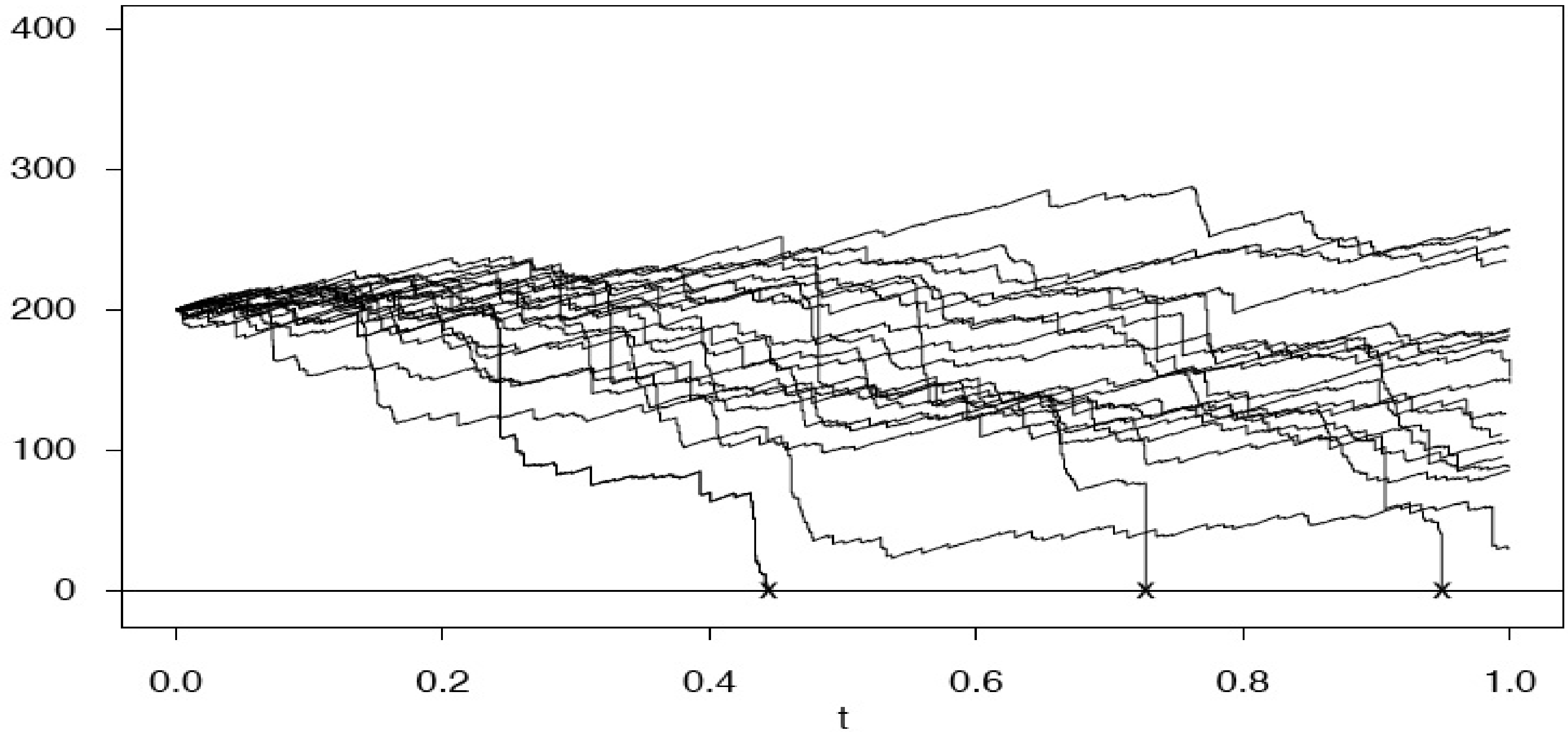
20 simulations from the ruin process C_t, $0 \leq t \leq 1$, with $(N(t)) \sim$ HPois$(100t)$ and $X_1 \sim$ Pareto$(2, 1)$.

Ruin Process with Exponential Claims



20 simulations from the ruin process C_t, $0 \leq t \leq 1$, with $(N(t))$ a doubly stochastic Poisson process with a two-state Markov intensity process: HPois$(10t)$ and HPois$(100t)$ with mean holding times 5 and 0.2, respectively, and $X_1 \sim$ Exp(1).

Ruin Process with Pareto Claims



20 simulations from the ruin process C_t, $0 \leq t \leq 1$, with $(N(t))$ a doubly stochastic Poisson process with a two-state Markov intensity process: HPois$(10t)$ and HPois$(100t)$ with mean holding times 5 and 0.2, respectively, and $X_1 \sim$ Pareto$(2, 1)$.

L3. The Capital Charge Problem within LDA

Loss Distribution Approach (cont'd)

Choose:

- Period T
- Distribution of $L_{i,k}^{T+1}$ for each cell i, k
- Interdependence between cells
- Confidence level $\alpha \in (0, 1)$, $\alpha \approx 1$
- Risk measure g_α

Capital charge for:

- Each cell: $C_{i,k}^{T+1,\text{OR}} = g_\alpha(L_{i,k}^{T+1})$
- Total OR loss: $C^{T+1,\text{OR}}$ based on $C_{i,k}^{T+1,\text{OR}}$

Basel II proposal

- **Period**: one year
- **Distribution**: should be based on
  - internal data/models
  - external data
  - expert opinion
- **Confidence level**: $\alpha = 99.9\%$, for economic capital purposes even $\alpha = 99.95\%$ or $\alpha = 99.97\%$
- **Risk measure**: VaR_α
- **Total capital charge**:

$$C^{T+1,\mathsf{OR}} = \sum_{i,k} \mathrm{VaR}_\alpha(L_{i,k}^{T+1})$$

  - possible reduction due to correlation effects

Basel II Proposal: Summary

- Marginal VaR calculations

$$\mathrm{VaR}_\alpha^1, \ldots, \mathrm{VaR}_\alpha^l$$

- Global VaR estimate

$$\mathrm{VaR}_\alpha^+ = \mathrm{VaR}_\alpha^1 + \cdots + \mathrm{VaR}_\alpha^l$$

- Reduction because of “correlation effects”

$$\mathrm{VaR}_\alpha < \mathrm{VaR}_\alpha^+$$

- Further possibilities: insurance, pooling, ...

Coherence and VaR

VaR_α is in general not coherent:

1. skewness

2. special dependence

3. very heavy-tailed losses

VaR_α is coherent for:

- elliptical distributions

Skewness

- 100 iid loans: 2%-coupon, 100 face value, 1% default probability (period: 1 year):

$$X_i = \begin{cases} -2 & \text{with probability 99\%} \\ 100 & \text{with probability 1\%} \quad \text{(loss)} \end{cases}$$

- Two portfolios $L_1 = \sum_{i=1}^{100} X_i$, $L_2 = 100X_1$

- $\underbrace{\mathrm{VaR}_{95\%}(L_1)}_{\mathrm{VaR}_{95\%}\left(\sum_{i=1}^{100} X_i\right)} > \underbrace{\mathrm{VaR}_{95\%}(100X_1)}_{\sum_{i=1}^{100} \mathrm{VaR}_{95\%}(X_i)}$ (!)

Special Dependence

- Given rvs $X_1, \ldots, X_n$ with marginal dfs $F_1, \ldots, F_n$, then one can always find a copula C so that for the joint model

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

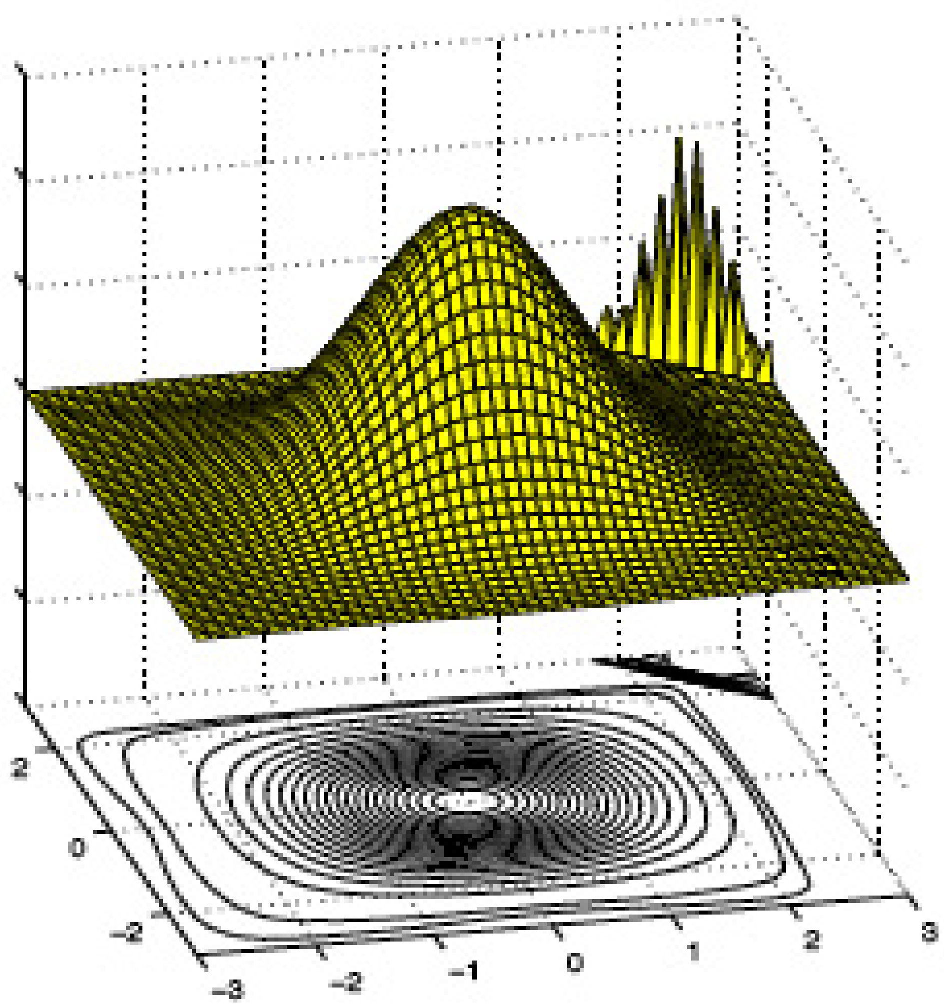
VaR_α is superadditive:

$$\mathrm{VaR}_\alpha\left(\sum_{k=1}^{n} X_k\right) > \sum_{k=1}^{n} \mathrm{VaR}_\alpha(X_k)$$

- In particular, take the “nice” case

$$F_1 = \cdots = F_n = N(0,1)$$

Special Dependence



Very Heavy-tailedness

- **Pareto:** take X_1, X_2 independent with $P(X_i > x) = x^{-1/2}$, $x \geq 1$ then for $x > 2$

$$P(X_1 + X_2 > x) = \frac{2\sqrt{x-1}}{x} > P(2X > x)$$

so that

$$\mathrm{VaR}_\alpha(X_1 + X_2) > \mathrm{VaR}_\alpha(2X_1) = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$$

- **Pareto-type:** similar result holds for X_1, X_2 independent with

$$P(X_i > x) = x^{-1/\xi} L(x),$$

where $\xi > 1$, L slowly varying

- For $\xi < 1$, we obtain subadditivity - WHY?

Several reasons:

- (Marcinkiewicz-Zygmund) Strong Law of Large Numbers

- Argument based on stable distributions

- Main reason however comes from functional analysis

  > *In the spaces $\mathcal{L}^p$, $0 < p < 1$, there exist no convex open sets other than the empty set and $\mathcal{L}^p$ itself.*

  Hence as a consequence 0 is the only continuous linear functional on $\mathcal{L}^p$; this is in violent contrast to $\mathcal{L}^p$, $p \geq 1$

- Discussion:
  - no reasonable risk measures exist
  - diversification goes the wrong way

Definition:
An $\mathbb{R}^d$-valued random vector $\boldsymbol{X}$ is said to be regularly varying if there exists a sequence (a_n), $0 < a_n \uparrow \infty$, $\mu \neq 0$ Radon measure on $\mathcal{B}\left(\overline{\mathbb{R}}^d \backslash \{0\}\right)$ with $\mu(\overline{\mathbb{R}}^d \backslash \mathbb{R}) = 0$, so that for $n \to \infty$,

$$nP(a_n^{-1}\boldsymbol{X} \in \cdot) \to \mu(\cdot) \quad \text{on } \mathcal{B}\left(\overline{\mathbb{R}}^d \backslash \{0\}\right).$$

Note that:

- $(a_n) \in RV_\xi$ for some $\xi > 0$
- $\mu(uB) = u^{-1/\xi}\mu(B)$ for $B \in \mathcal{B}\left(\overline{\mathbb{R}}^d \backslash \{0\}\right)$

Theorem: (several versions – Samorodnitsky)
If $(X_1, X_2)' \in RV_{-1/\xi}$, $\xi < 1$, then for α sufficiently close to 1, VaR_α is subadditive.

Phase Transition of Value-at-Risk

Theorem:

Assume that $\boldsymbol{X} = (X_1, X_2)$ is a rv such that

- $X_i \sim F$ for all i where F is continuous and $\overline{F} \in RV_{-\beta}$, $\beta > 0$
- $-\boldsymbol{X}$ has an Archimedean copula with generator ψ which is regularly varying at 0 with index $-\delta < 0$

then there exists a constant $q_2(\delta, \beta)$such that

$$\mathrm{VaR}_\alpha(X_1 + X_2) \sim (q_2(\delta, \beta))^{1/\beta} \mathrm{VaR}_\alpha(X_1), \quad \alpha \to 1$$

- Behavior of $q_2(\delta, \beta)$ with respect to β and δ, respectively

L4. Marginal VaR Estimation

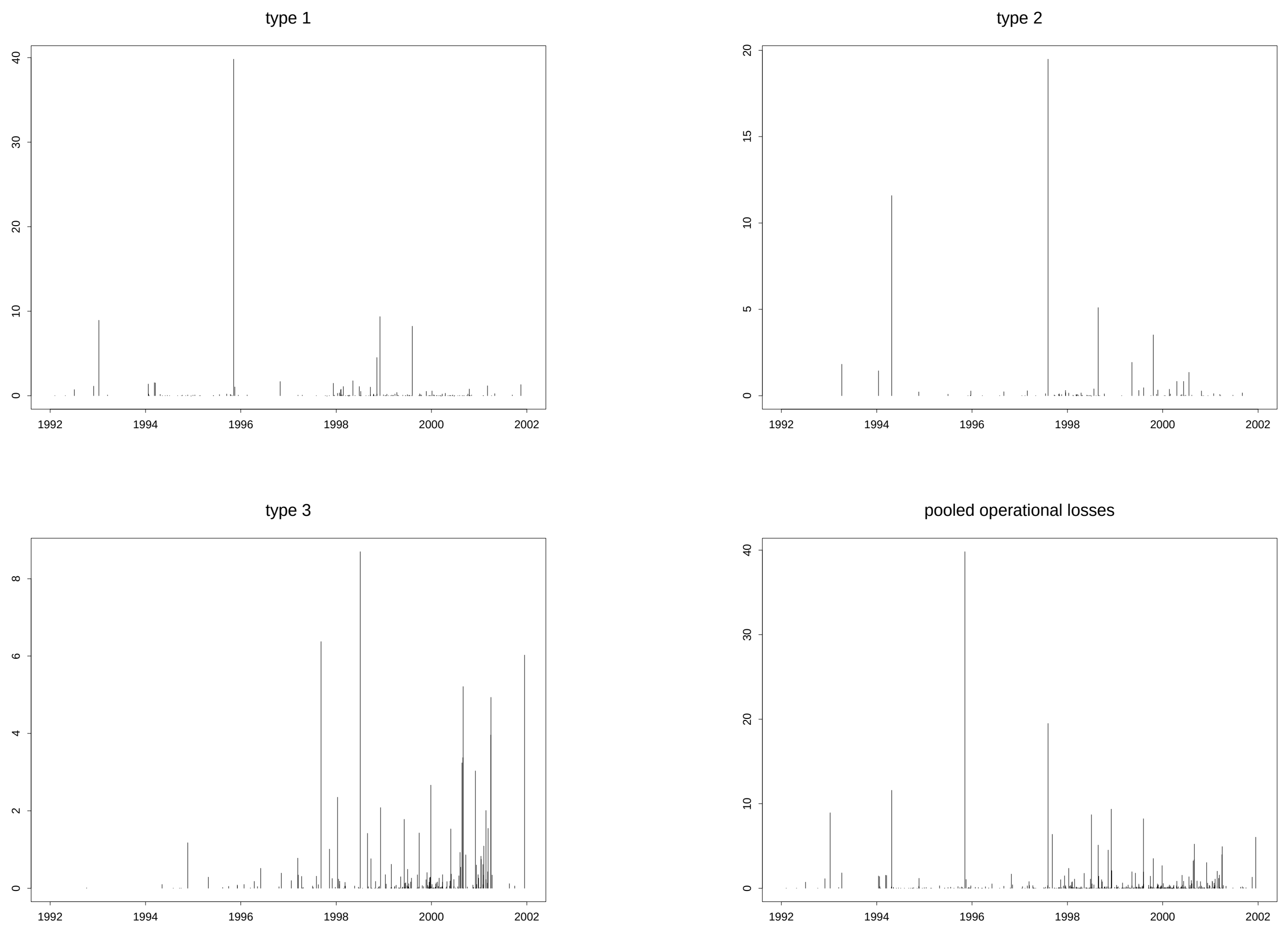
LDA revisited

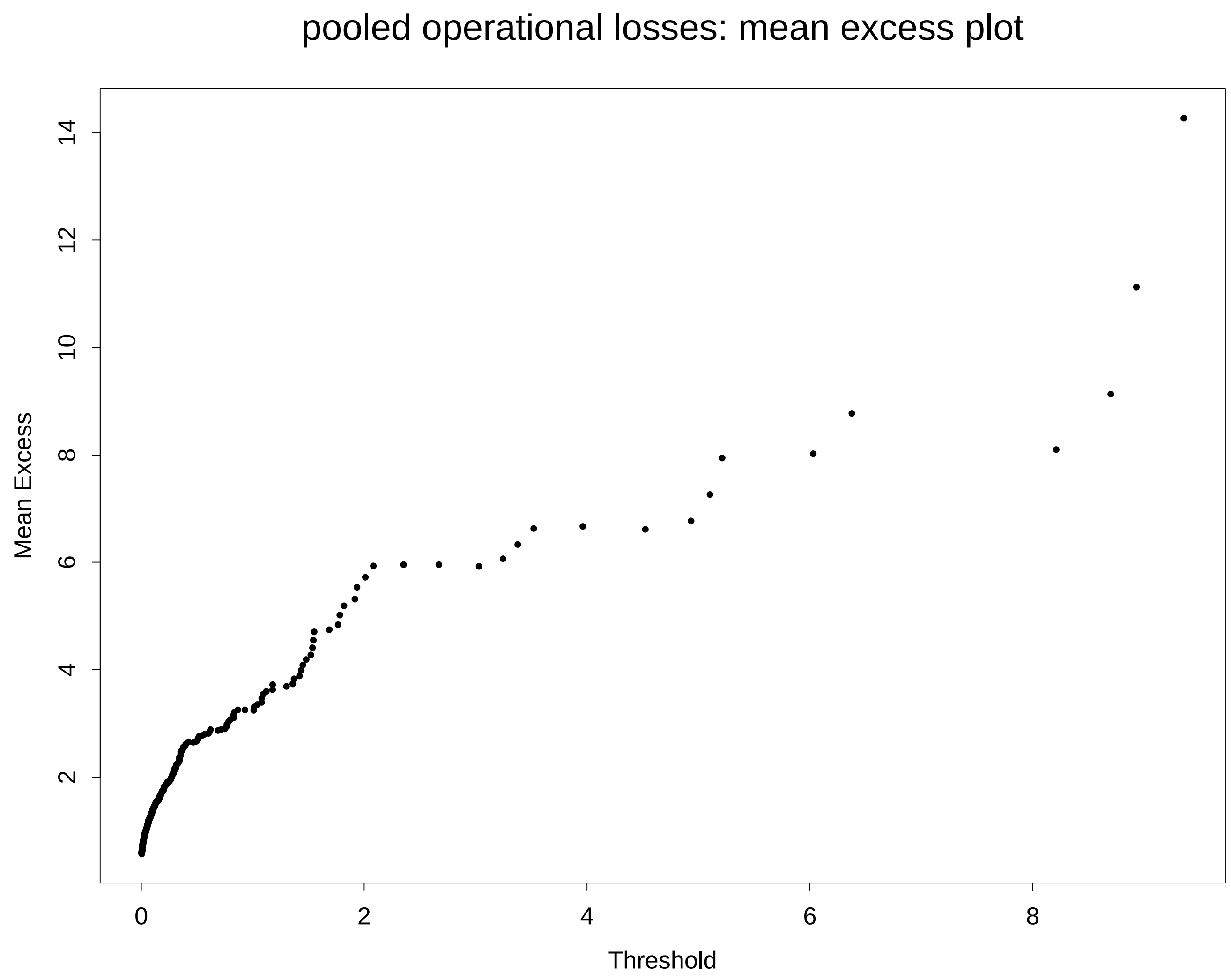
- **Recall:** $\mathrm{VaR}_\alpha^{i,k}$ is a Value-at-Risk of a compound sum

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

- **Tasks:**
  - Suitable model for the severity $X_{i,k}$
  - Suitable model for the frequency $N_{i,k}^{T+1}$
  - Estimation of $\mathrm{VaR}_\alpha^{i,k}$

Some OpRisk Data





- $P(L > x) \sim x^{-1/\xi} L(x)$

Stylized Facts

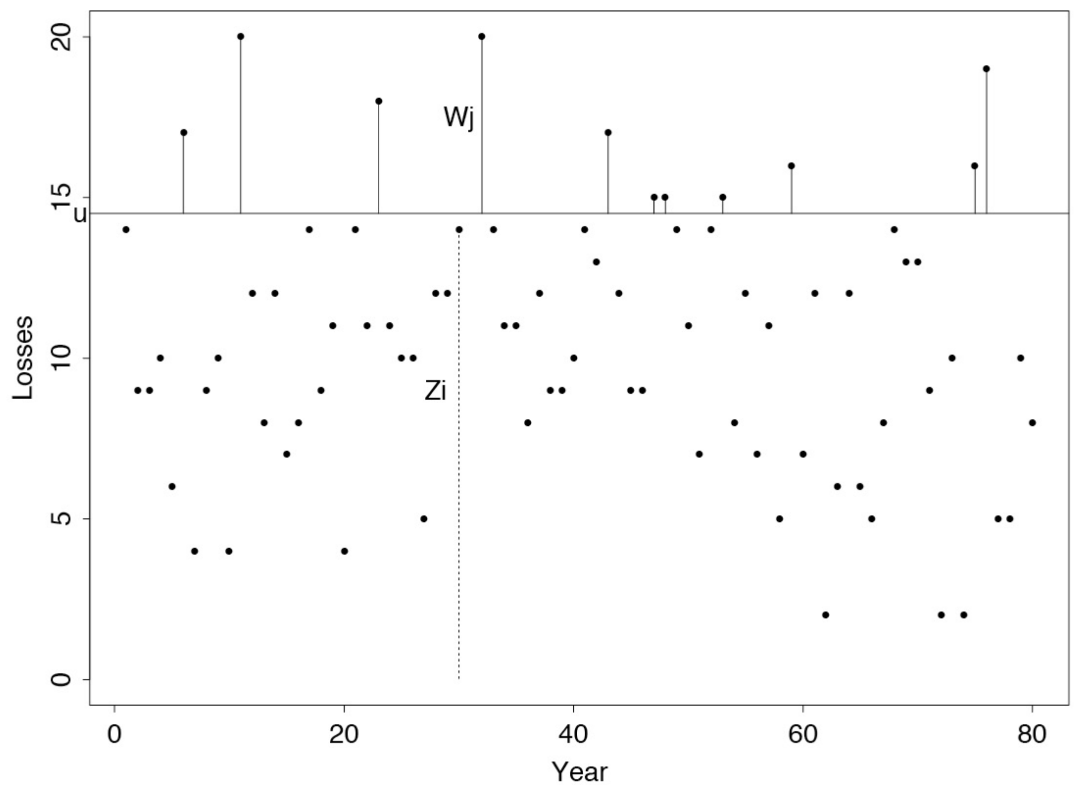
- Stylized facts about OpRisk losses:
  - Loss amounts show extremes
  - Loss occurence times are irregularly spaced in time (reporting bias, economic cycles, regulation, management interactions, structural changes, . . .)
  - Non-stationarity (frequency(!), severity(?))

- Large losses are of main concern

- Repetitive versus non-repetitive losses

- **However**: severity is of key importance

Peaks-over-Threshold Method



- Distribution of the exceedances
- Distribution of the inter-arrival times

Peaks-over-Threshold Method (POT)

$X_1, \ldots, X_n$ iid with distribution function F satisfying

$$\overline{F}(x) = x^{-1/\xi} L(x), \quad \xi > 0 \text{ and } L \text{ slowly varying}$$

- **Excess distribution:** asymptotically Generalized Pareto (GPD)

$$P(X - u > y | X > u) \sim \left(1 + \xi \frac{y}{\beta(u)}\right)^{-1/\xi}, \quad u \to \infty$$

➡ **POT-MLE estimation** of tail probabilities and risk measures

$$\widehat{\overline{F}}(x) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x-u}{\hat{\beta}}\right)^{-1/\hat{\xi}}, \quad x > u$$

$$\widehat{\mathrm{VaR}}_\alpha = u + \frac{\hat{\beta}}{\hat{\xi}} \Big(\big(\frac{n}{N_u} (1-\alpha) \big)^{-\hat{\xi}} - 1 \Big), \quad \alpha \text{ “close” to } 1$$

Threshold Choice

Application of POT based estimates requires a choice of an appropriate threshold u

- **Rates of convergence:** very tricky
  - No generally valid convergence rate
  - Convergence rate depends on F, in particular on the slowly varying function L, in a complicated way and may be very slow
  - L is not visible from data directly

➡ Threshold choice is very difficult. Trade-off between bias and variance usually takes place.

- **Diagnostic tools:**
  - Graphical tools (mean excess plot, shape plot,...)
  - Bootstrap and other methods requiring extra conditions on L

Basel II QIS 2002 Data

- **POT Analysis of Severities**: $P(L_i > x) = x^{-1/\xi_i} L_i(x)$

Business line	$\widehat{\xi}_i$
Corporate finance	1.19 (*)
Trading & sales	1.17
Retail banking	1.01
Commercial banking	1.39 (*)
Payment & settlement	1.23
Agency services	1.22 (*)
Asset management	0.85
Retail brokerage	0.98
	* means significant at 95% level

$\widehat{\xi}_i > 1$: infinite mean

- **Remark**: different picture at level of individual banks

Issues Regarding Infinite Mean Models

- Reason for $\xi > 1$?
- Potentially:
  - wrong analysis
  - EVT conditions not fulfilled
  - contamination, mixtures
- We concentrate on the latter:
  Two examples:

  ❶ Contamination above a high threshold
  ❷ Mixture models

- Main aim: show through examples how certain data-structures can lead to infinite mean models

Contamination Above a High Threshold

Example 1: Consider the model

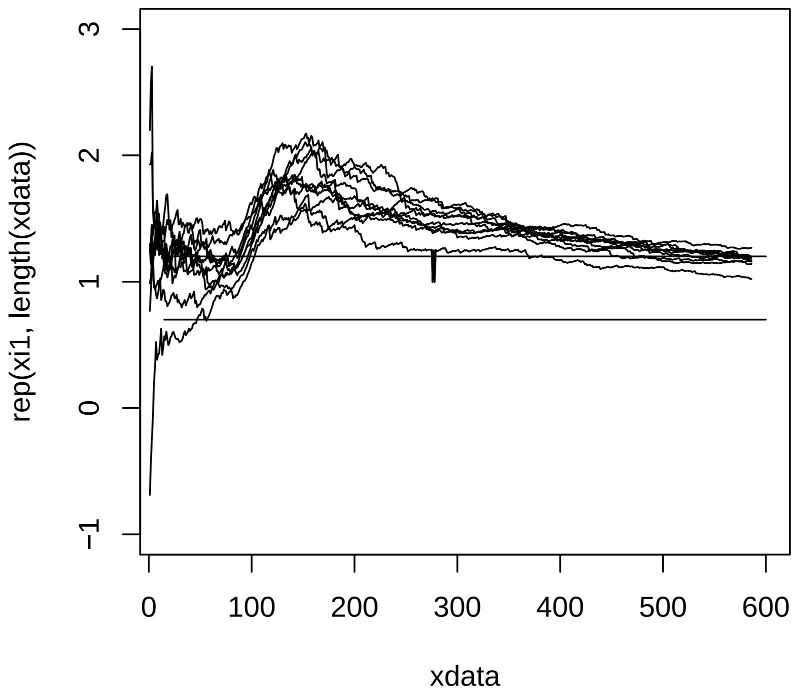
$$F_X(x) = \begin{cases} 1 - \left(1 + \frac{\xi_1 x}{\beta_1}\right)^{-1/\xi_1} & \text{if} \quad x \le v, \\ 1 - \left(1 + \frac{\xi_2 (x - v^*)}{\beta_2}\right)^{-1/\xi_2} & \text{if} \quad x > v, \end{cases}$$

where $0 < \xi_1 < \xi_2$ and $\beta_1, \beta_2 > 0$.

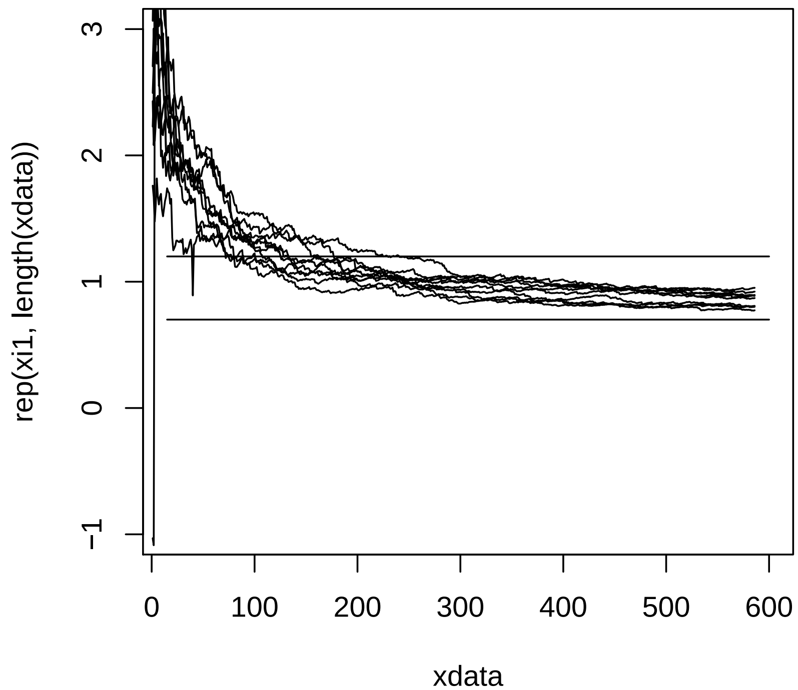
- v^* is a constant depending on the model parameters in a way that F_X is continuous
- VaR can be calculated explicitly:

$$\mathrm{VaR}_\alpha(X) = \begin{cases} \frac{1}{\xi_1}\beta_1 \left((1-\alpha)^{-\xi_1} - 1\right) & \text{if} \quad \alpha \le F_X(v), \\ v^* + \frac{1}{\xi_2}\beta_2 \left((1-\alpha)^{-\xi_2} - 1\right) & \text{if} \quad \alpha > F_X(v). \end{cases}$$

Shape Plots

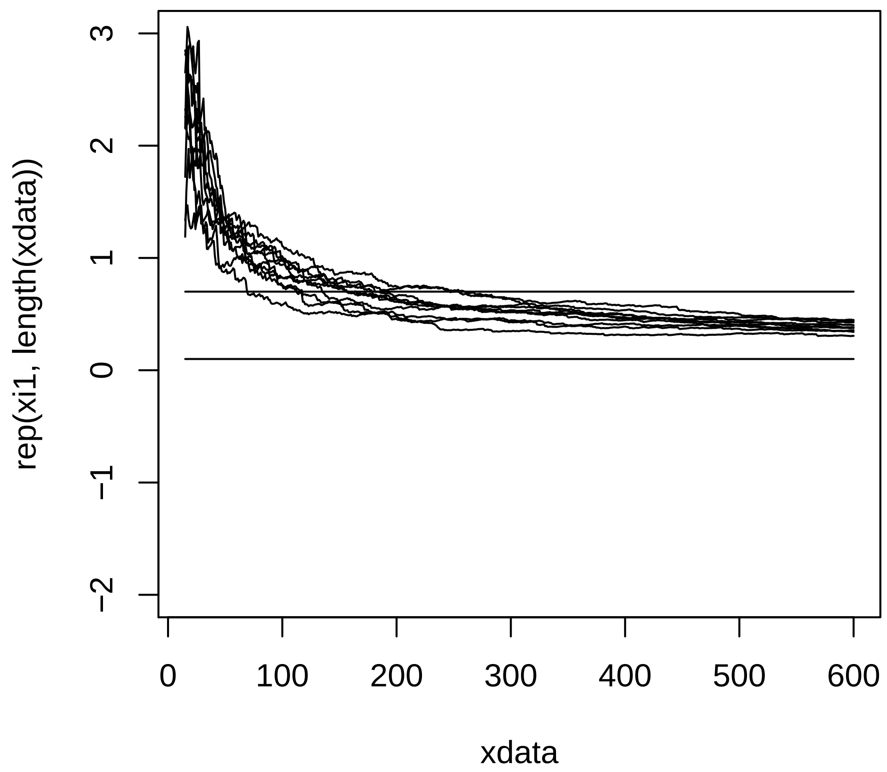


Easy case: v low



Hard case: v high

Finite Mean Case



Careful: similar picture for v low and $\xi_1 \ll \xi_2 < 1$

Contamination above a high threshold (cont'd)

- **Easy case: v low**
  - Change of behavior typically visible in the mean excess plot

- **Hard case: v high**
  - Typically only few observations above v
  - Mean excess plot may not reveal anything
  - Classical POT analysis easily yields incorrect resuls
  - Vast overestimation of VaR possible

Mixture Models

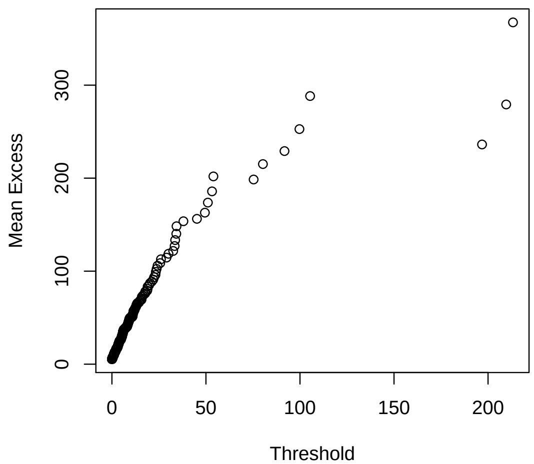
Example 2: Consider

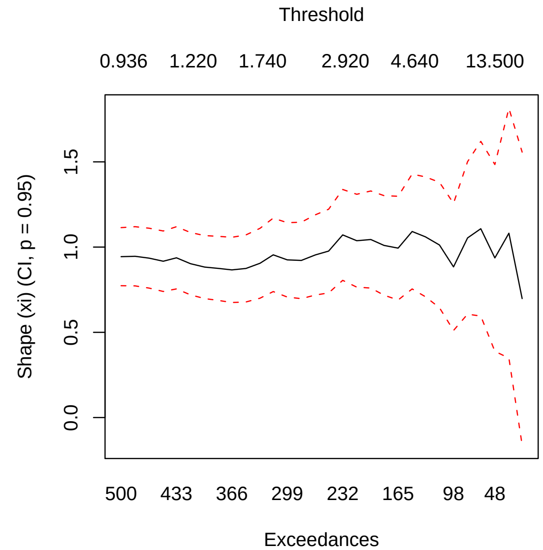
$$F_X = (1-p)F_1 + pF_2,$$

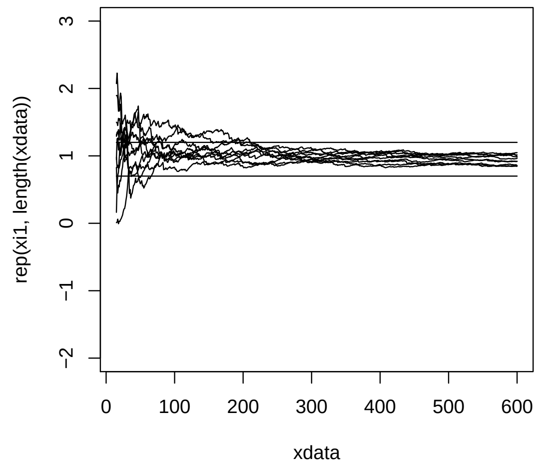
with F_i exact Pareto, i.e. $F_i(x) = 1 - x^{-1/\xi_i}$ for $x \geq 1$ and $0 < \xi_1 < \xi_2$.

- Asymptotically, the tail index of F_X is ξ_2
- VaR_α can be obtained numerically and furthermore
  - does not correspond to VaR_α of a Pareto distribution with tail-index ξ^*
  - equals VaR_{α^*} corresponding to F_2 at a level α^* lower than α

- Classical POT analysis can be very misleading:







VaR for Mixture Models

α	$\mathrm{VaR}_\alpha(F_X)$	$\mathrm{VaR}_\alpha(\text{Pareto}(\xi_2))$	ξ^*
0.9	6.39	46.42	0.8
0.95	12.06	147.36	0.83
0.99	71.48	2154.43	0.93
0.999	2222.77	10^5	1.12
0.9999	10^5	$4.64 \cdot 10^6$	1.27
0.99999	$4.64 \cdot 10^6$	$2.15 \cdot 10^8$	1.33

Value-at-Risk for mixture models with $p = 0.1$, $\xi_1 = 0.7$ and $\xi_2 = 1.6$.

Including Frequencies

The POT method can be embedded into a wider framework based on Point processes

- **iid case:** exceedance times follow asymptotically a homogeneous Poisson Process

- **Extensions:** several possibilities
  - Including the severities: marked Poisson process
  - Non-stationarity: non-homogeneous Poisson processes
  - Over-dispersion: doubly stochastic processes
  - Short-range dependence: clustering

VaR Estimation for Compound Sums

Proposition:

Let $X, X_1, X_2, \ldots$ be iid with $F_X \in \mathcal{S}$. If for some $\varepsilon > 0$, $\sum_{n=1}^{\infty}(1+\varepsilon)^n P(N=n) < \infty$ (satisfied for instance in the important binomial, Poisson and negative binomial cases), then

$$\lim_{x\to\infty} \frac{P(\sum_{i=1}^{N} X_i > x)}{1 - F_X(x)} = E(N).$$

Approximation of VaR:

$$\mathrm{VaR}_\alpha\left(\sum_{i=1}^{N} X_i\right) \sim \mathrm{VaR}_{\alpha^*}(X), \quad \alpha^* = 1 - \frac{1-\alpha}{EN}, \quad \alpha \to 1$$

Marginal VaR Estimation

Approximative method

1. Estimate the excess distribution of severities using POT

2. Calculate

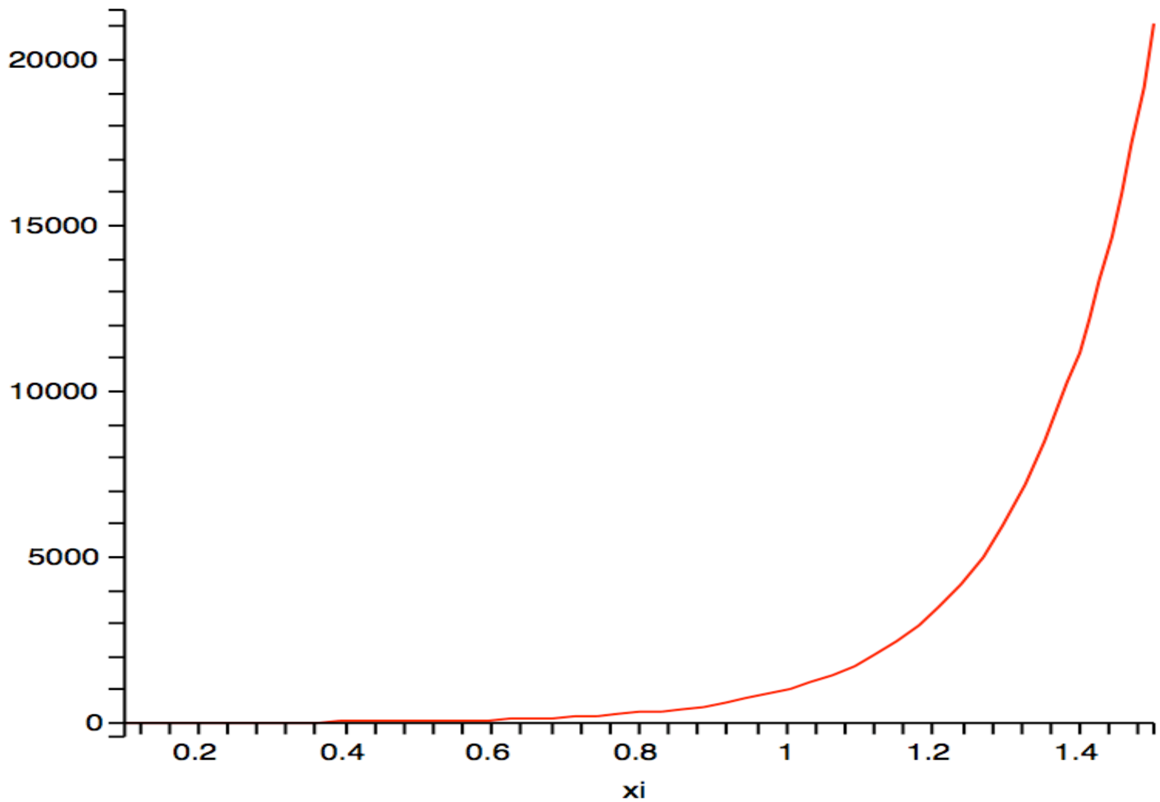
$$\widehat{\mathrm{VaR}}_\alpha(L_{i,k}^{T+1}) = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\left(\frac{n(1-\alpha)}{N_u E(N_i^{T+1})} \right)^{-\hat{\xi}} - 1 \right)$$

Monte Carlo methods

1. Choose a distribution for severities and a process for frequencies (evt. jointly)

2. After a large number of simulations, estimate the VaR of the compound sum via the POT-MLE method

POT-MLE VaR Estimate



$\mathrm{VaR}_{99.9\%}$ of a GPD$(\xi, 1)$ rv. as a function of $\xi \in (0, 1.5)$

- Small changes in ξ lead to considerable changes in VaR

Marginal VaR Estimate: Issues

- $\widehat{\mathrm{VaR}}_\alpha$ is an exponential function of ξ and therefore very sensitive to $\hat{\xi}$

- Confidence intervals for $\widehat{\mathrm{VaR}}_\alpha$ widen rapidly with increasing α and decreasing sample size

- **Fazit 1:** for very high levels (99.9% or 99.97%) there is typically substantial uncertainty and variability in the VaR estimate due to the lack of data

- **Fazit 2:** Issues far in the tail call for judgement

- $\hat{\xi} > 1$ **is an issue!**

L5. Global VaR Estimation

Recall:

- Global VaR estimate

$$\mathrm{VaR}_\alpha^+ = \mathrm{VaR}_\alpha^1 + \cdots + \mathrm{VaR}_\alpha^l$$

- Reduction because of "correlation effects"

$$\mathrm{VaR}_\alpha < \mathrm{VaR}_\alpha^+$$

- **In general, VaR_α^+ is not the upper bound!**

Bounds on VaR

Find optimal bounds for

$$\mathrm{VaR}_{l,\alpha}^{T+1} \leq \mathrm{VaR}_{\alpha}^{T+1}\left(\sum_{k=1}^{d} L_k^{T+1}\right) \leq \mathrm{VaR}_{u,\alpha}^{T+1}$$

given marginal VaR's and dependence information

Solution:

- Fréchet Problem
- Mass Transportation Problem

Example 1: Comonotonic Case

Recall:

$L_1^{T+1}, \ldots, L_d^{T+1}$ are comonotonic if there exists a rv Z and increasing functions $f_1^{T+1}, \ldots, f_d^{T+1}$, so that

$$L_i^{T+1} = f_i^{T+1}(Z), \quad i = 1, \ldots, d$$

- If $L_1^{T+1}, \ldots, L_d^{T+1}$ are comonotonic VaR is additive

$$\mathrm{VaR}_\alpha^{T+1}\left(\sum_{k=1}^d L_k^{T+1}\right) = \sum_{k=1}^d \mathrm{VaR}_{k,\alpha}^{T+1}$$

Example 2: No Dependence Information

Take $L_i^{T+1} = L_i, \;\; i = 1, \ldots, d = 8$ and

$$F_{L_1} = \cdots = F_{L_d} = \text{Pareto}(1, 1.5)$$

- **Comonotonic case:**

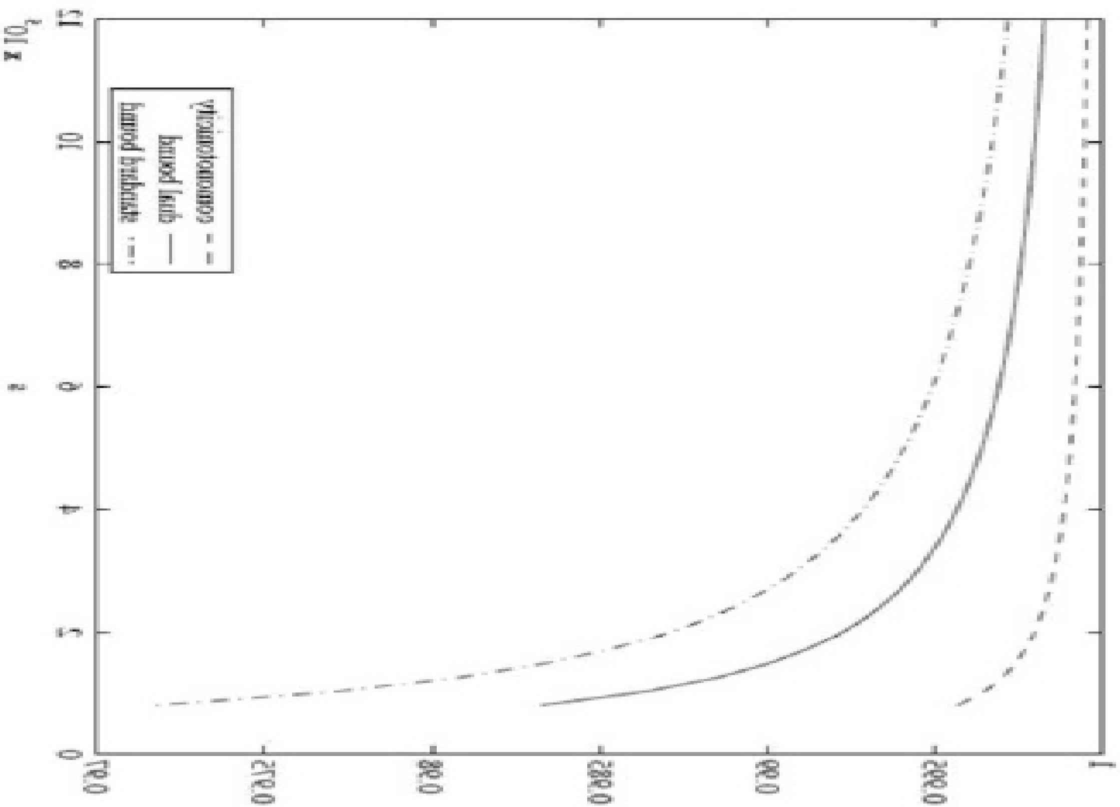
$$VaR_{0.999}\left(\sum_{i=1}^{8} L_i\right) = \sum_{i=1}^{8} \mathrm{VaR}_{0.999}(L_i) = 0.79$$

- **Unconstrained upper bound:**

$$\mathrm{VaR}_{0.999}\left(\sum_{i=1}^{8} L_i\right) \leq 1.93$$

Example 3: No Dependence Information

$F_{L_1} \neq \cdots \neq F_{L_d}$ is more difficult



Bounds on VaR using the OpRisk portfolio given in Moscadelli(2004)

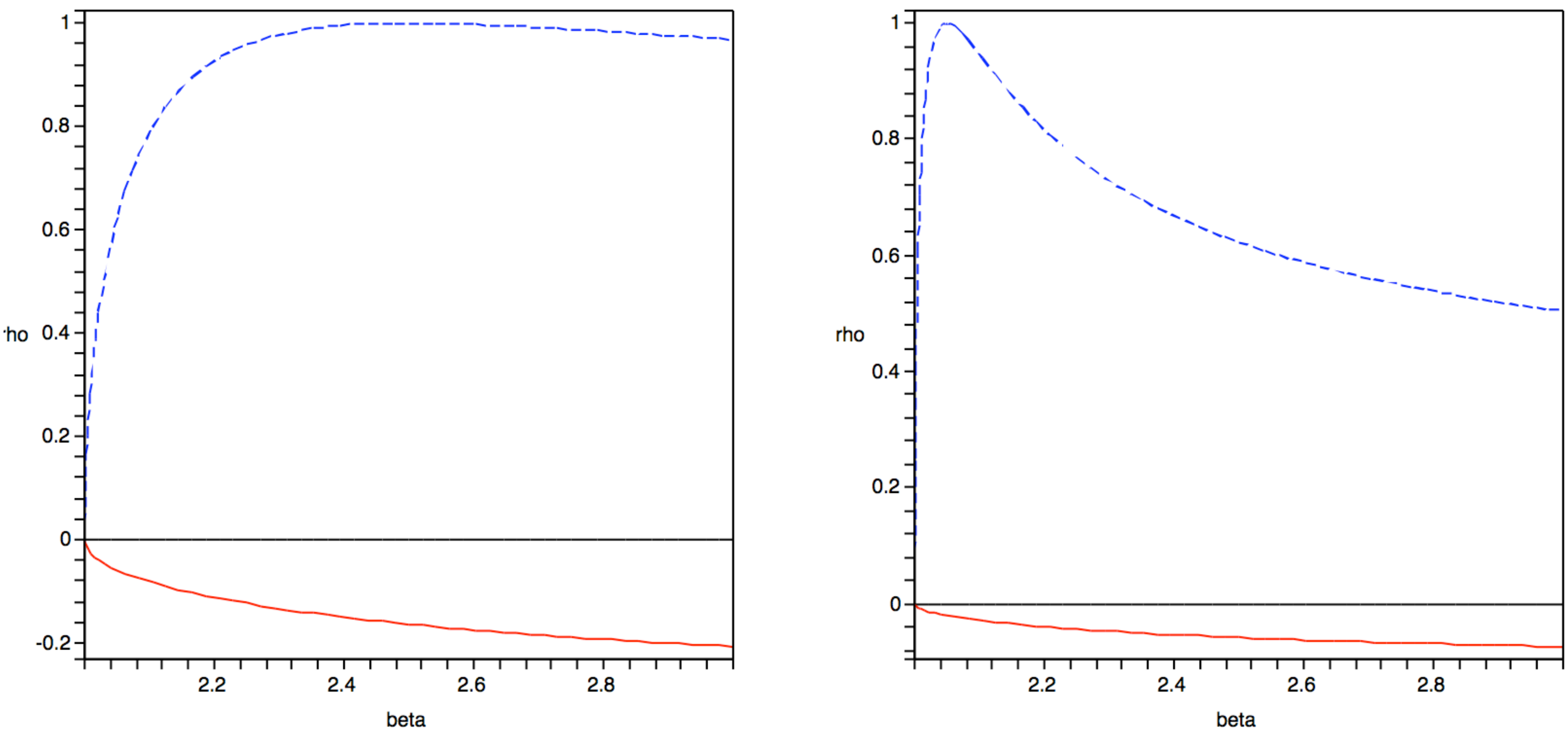
Correlation

Correlation:

Correlation (linear, rank, tail) is one-number summary: ρ, τ, ρ_S ...

- Careful: linear correlation does not exist for $\xi > 0.5$
- Linear correlation is typically small for heavy tailed rvs
- Knowledge of correlation (linear, rank, tail...) is sufficient for individual models, but totally insufficient in general

Correlation



Upper and lower bound on linear correlation $\rho(L_1, L_2)$ for $L_1 \sim$ Pareto(2.5) (left) and $L_1 \sim$ Pareto(2.05) (right) and $L_2 \sim$ Pareto(β)

Copulas

Copula:

With $L_i^{T+1} \sim F_i$ the joint distribution can be written as

$$P(L_1^{T+1} \leq l_1, \ldots, L_d^{T+1} \leq l_d) = C(F_1(l_1), \ldots, F_d(l_d))$$

The function C is known as copula and is a joint distribution on $[0,1]^d$ with uniform marginals

- A copula and marginal distributions determine the joint model completely
- However: there are not enough OpRisk data: one year of loss data comprises to a single observation of $(L_1^{T+1}, \ldots, L_d^{T+1})$

Dynamic Dependence Models

In order to use the data at hand, we need a dynamic model for the compound processes.

Consider:

$$d = 2: \quad L_k = \sum_{i=1}^{N_k(T)} X_{i,k}, \quad k = 1, 2$$

and

1. make $(X_{i,1})$ and $(X_{i,2})$ dependent, $i \geq 1$
2. make $\{N_1(t) : t \leq T\}$ and $\{N_2(t) : t \leq T\}$ dependent
3. combination of both

Dependent Counting Processes

- Various models for dependent counting processes $\{N_1(t) : t \leq T\}$ and $\{N_2(t) : t \leq T\}$ exist:
  - Common shock models
  - Point process models
  - Mixed Poisson processes with dependent mixing rvs.
  - Lévy Copulas
- So far, there is no general dependence concept
- It is not clear how to quantify dependence between processes
- It is less transparent how the dependence structure of the frequency processes affects the dependence structure of the compound rvs

One Loss Causes Ruin Problem

Question: how do the marginal severities affect the global loss?

- based on Lorenz curve in economics
  - 20 – 80 rule for $1/\xi = 1.4$
  - 0.1 – 95 rule for $1/\xi = 1.01$

Proposition 1

For $L_1, \ldots, L_d$ iid and subexponential we have for

$L = L_1 + \cdots + L_d$ that

$$P(L > x) \sim dP(L_1 > x), \quad x \to \infty$$

Proposition 2

Suppose in addition that $L_i = \sum_{k=1}^{N_i} X_{i,k}$ where $N_i \sim \text{Poi}(\lambda_i)$ are independent. Furthermore, for $i = 1, \ldots, d$, $X_{i,k}$'s are iid with

$$P(X_i > x) = x^{-1/\xi_i} h_i(x), \quad h_i \text{ slowly varying}$$

If $\xi_1 > \cdots > \xi_d$, we have that

$$P(L > x) \sim cP(X_1 > x)$$

Discussion

- **Tail issues**:
  - robust statistics
  - scaling
  - mixtures
- **Infinite mean**: industry occasionally uses
  - constrained estimation to $\xi < 1$
  - estimate under the condition of a finite upper limit

Discussion (cont'd)

- **Aggregation issues**:
  - adding risk measures across a 7×8 table
  - reduction because of "correlation effects"

- **Data issues**:
  - impact of pooling
  - incorporation of external data and expert opinion
  - credibility theory

References

[1] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.

[2] McNeil, A.J., Frey, R., and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press.

[3] Moscadelli, M. (2004) The Modelling of Operational Risk: Experience with the Analysis of the Data, Collected by the Basel Committee, Banaca d'Italia, report 517-July 2004

[4] Nešlehová, J., Embrechts, P. and Chavez-Demoulin, V. (2006) Infinite mean models and the LDA for operational risk. Journal of Operational Risk, 1(1), 3-25.