Quantitative Models for Operational Risk: Extremes, Dependence and Aggregation

Valérie Chavez-Demoulin

Paul Embrechts

Johanna Nešlehová

ETH Zurich

www.math.ethz.ch/~embrechts

valerie@math.ethz.ch

embrechts@math.ethz.ch

johanna@math.ethz.ch

Introduction

Context:

- Basel II (Banking) and Solvency 2 (Insurance)
- AMA approach to Operational Risk

Our contribution:

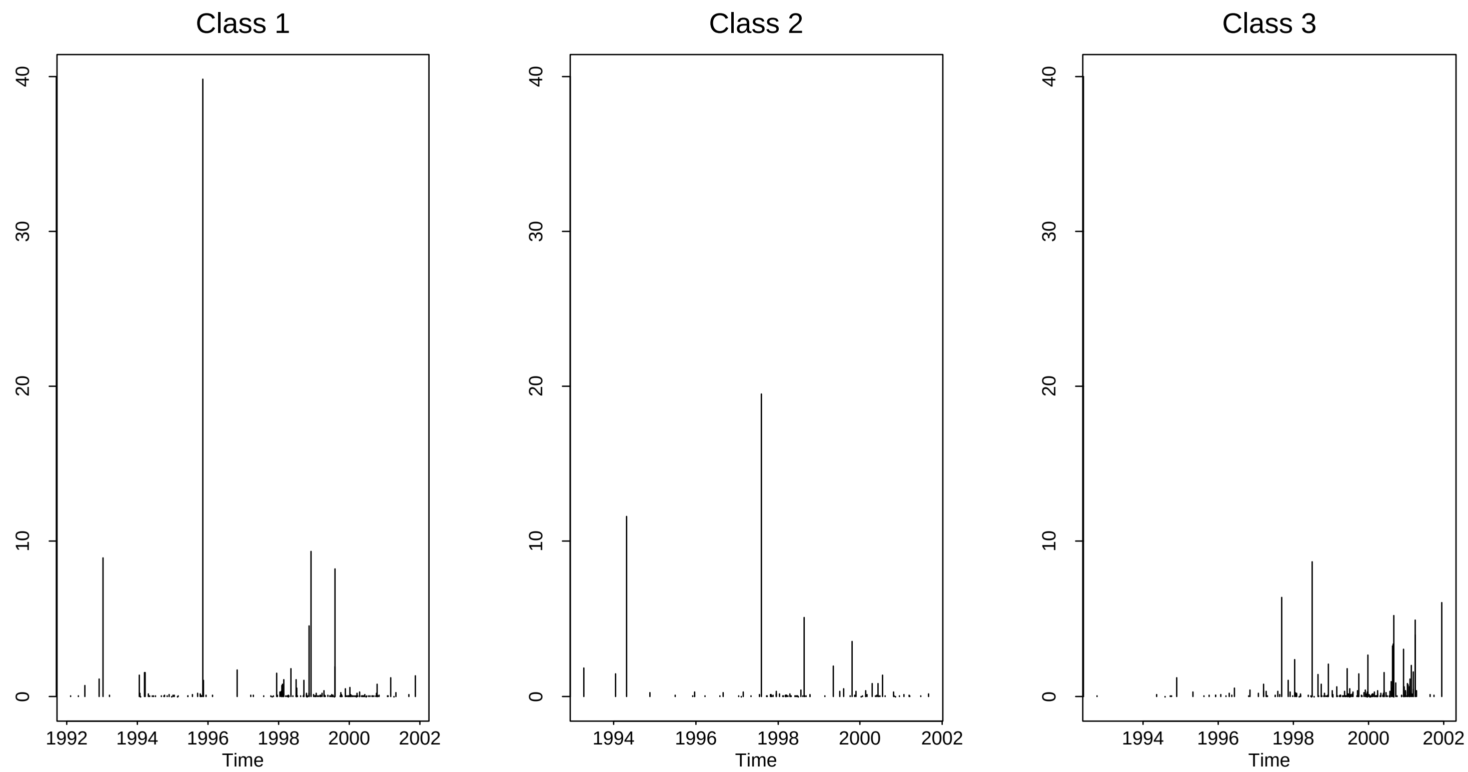
- Based on stylized facts of OpRisk data
- New stochastic methodology
- No final AMA solution, but contribution towards such a solution

Outline of the talk

- Stylized facts of OpRisk data
- Methodological issues
- Non-stationary EVT models
- Dependence and Point Processes
- Aggregation
- Conclusion

I. Stylized facts on OpRisk data

Some operational loss data:



Stylized facts

- Different, non-homogeneous classes
- Extremes matter
- Non-repetitive/large claims versus repetitive/small claims
- Non-stationarity
- Interdependence
- Restricted data warehouses
- Thresholding

QIS-3: above facts are corroborated (Moscadelli)

Regulatory issues

- An appropriate (?) risk measure:
  - Value-at-Risk
  - 1 year
  - $\alpha \geq 0.99$ (even 0.9997)

- Appropriate (?)
  - (non-) coherence ?
  - beyond VaR: expected shortfall ?
  - (very) extreme quantiles
  - yearly data: hardly any (scaling ?)

II. Methodological issues

- EVT matters:
  - very high quantiles
  - non-stationarity

- Dependence matters:
  - "correlating" loss processes
  - copulas as a method for going beyond correlation
  - how to quantify dependence ?

- Aggregation matters:
  - global 1-year OpRisk measure
  - restricted/partial information
  - aggregate risk measures across risk classes

III. Non-stationary EVT models

Basic EVT methodology

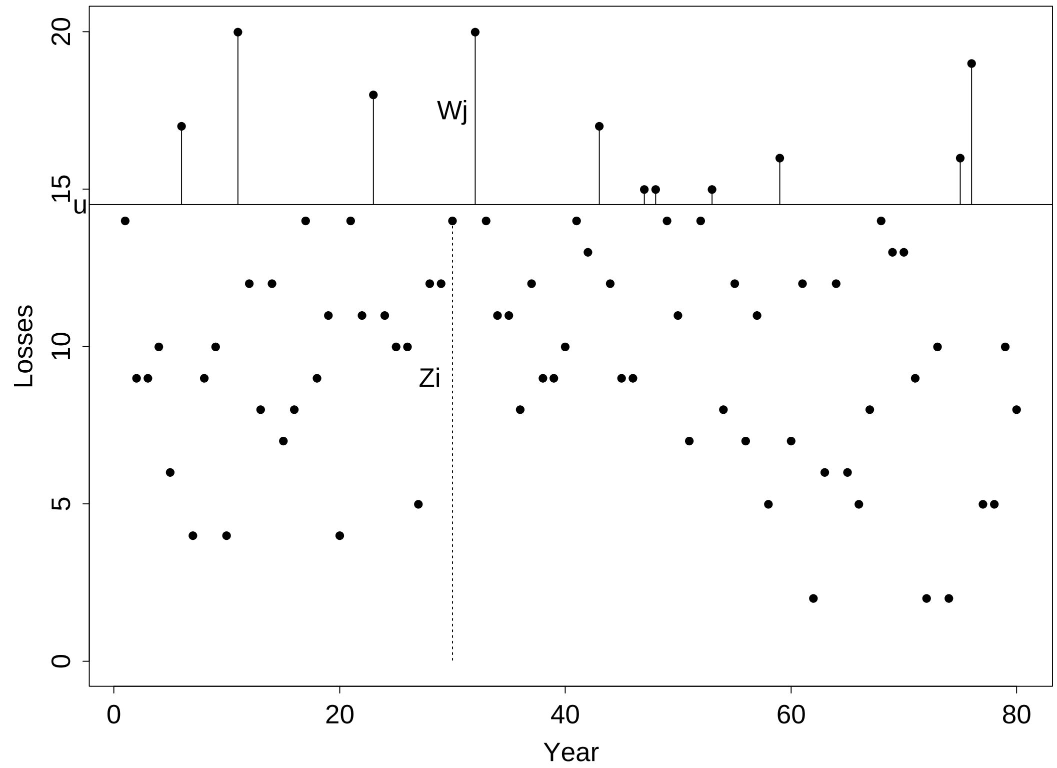
- Ground-up losses are denoted by $Z_1, Z_2, \ldots, Z_q$

- u is a typically high threshold

- $W_1, \ldots, W_n$ are the excess losses from $Z_1, \ldots, Z_q$ above u, i.e.

$$W_j = Z_i - u \quad \text{for some } j = 1, \ldots, n \text{ and } i = 1, \ldots, q, \text{ where } Z_i > u$$

The Point Process of Exceedances within the Peaks over Threshold (POT) Method



The Peaks Over Threshold Method

For iid losses, and for u high enough:

- Conditional losses $W_1, \ldots, W_n$ follow a Generalized Pareto Distribution (GPD)

$$G_{\kappa,\sigma}(w) = \begin{cases} 1 - (1 + \kappa w/\sigma)_+^{-1/\kappa}, & \kappa \neq 0 \\ 1 - \exp(-w/\sigma)\,, & \kappa = 0 \end{cases}$$

- The exceedance points of $Z_1, \ldots, Z_q$ of the threshold u follow (approximately) a homogeneous Poisson process with intensity $\lambda > 0$

- The conditional losses $W_1, \ldots, W_n$ and the Poisson exceedance process are (approximately) independent, and

- An approximate log-likelihood function $l(\lambda, \sigma, \kappa)$ can be derived

An advanced POT model: non-stationarity

In general:

$$(\lambda, \kappa, \sigma) \quad \Rightarrow \quad (\lambda(t,\tau), \kappa(t,\tau), \sigma(t,\tau))$$

For pooled OpRisk:

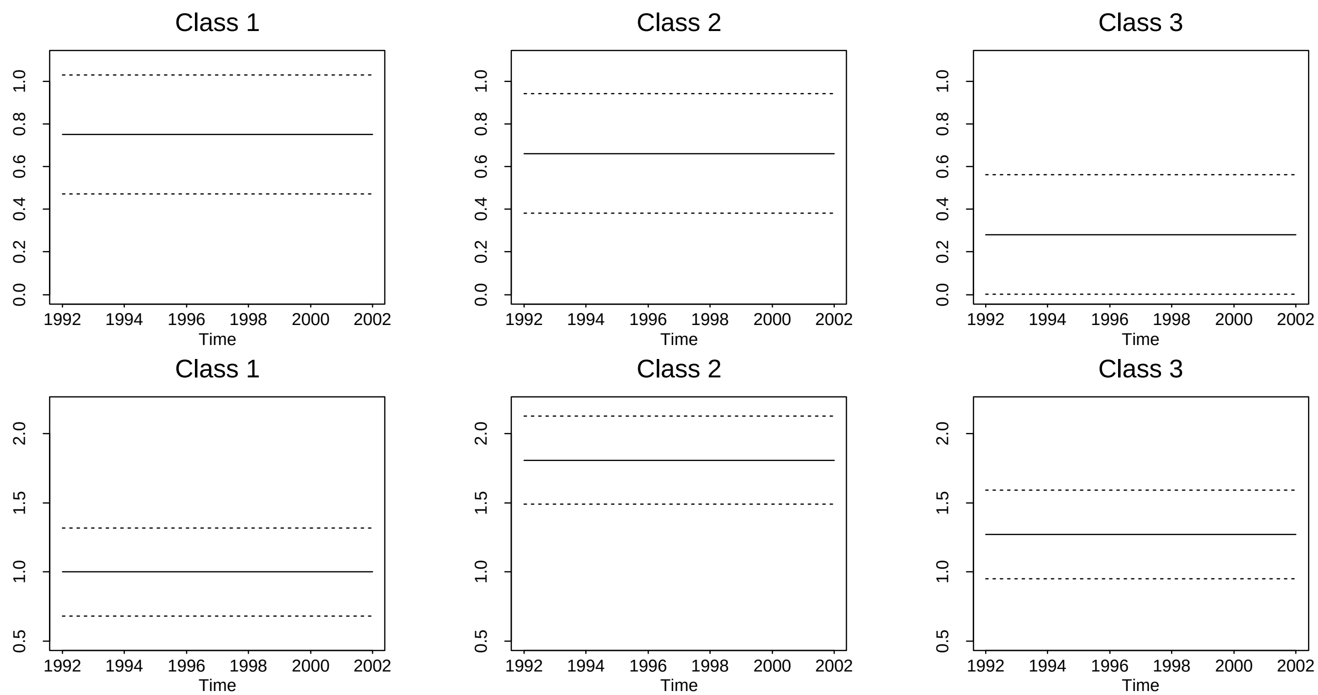
$$\log \lambda(t,\tau) = \gamma_\tau I_\tau + \beta I_{(t>t_c)} + g(t)$$

where

- I_τ class indicator
- $I_{(t>t_c)}$ change point indicator
- $g(t)$ general smooth function

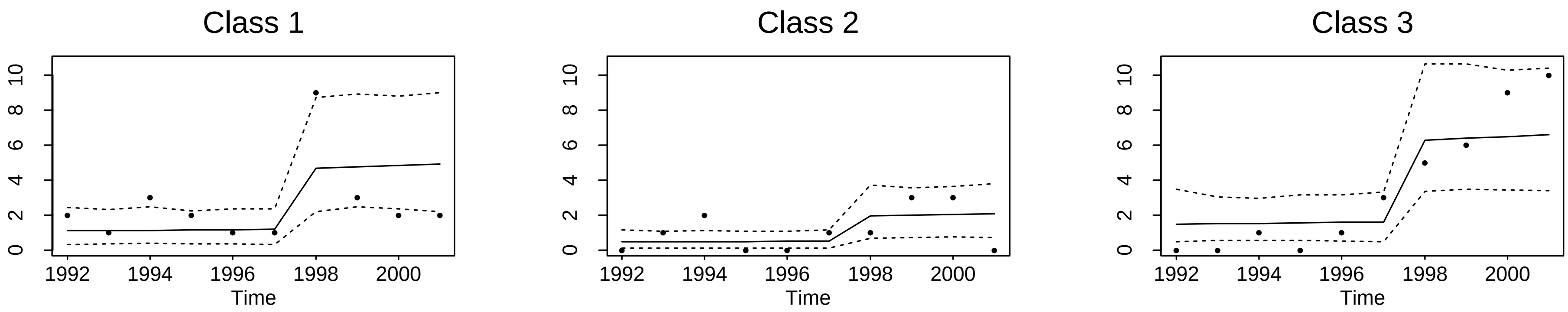
The advanced POT model: results of fit to pooled OpRisk data

- $\hat{\kappa} = \hat{\kappa}(\tau)$ (upper panels)
- $\hat{\sigma} = \hat{\sigma}(\tau)$ (lower panels)

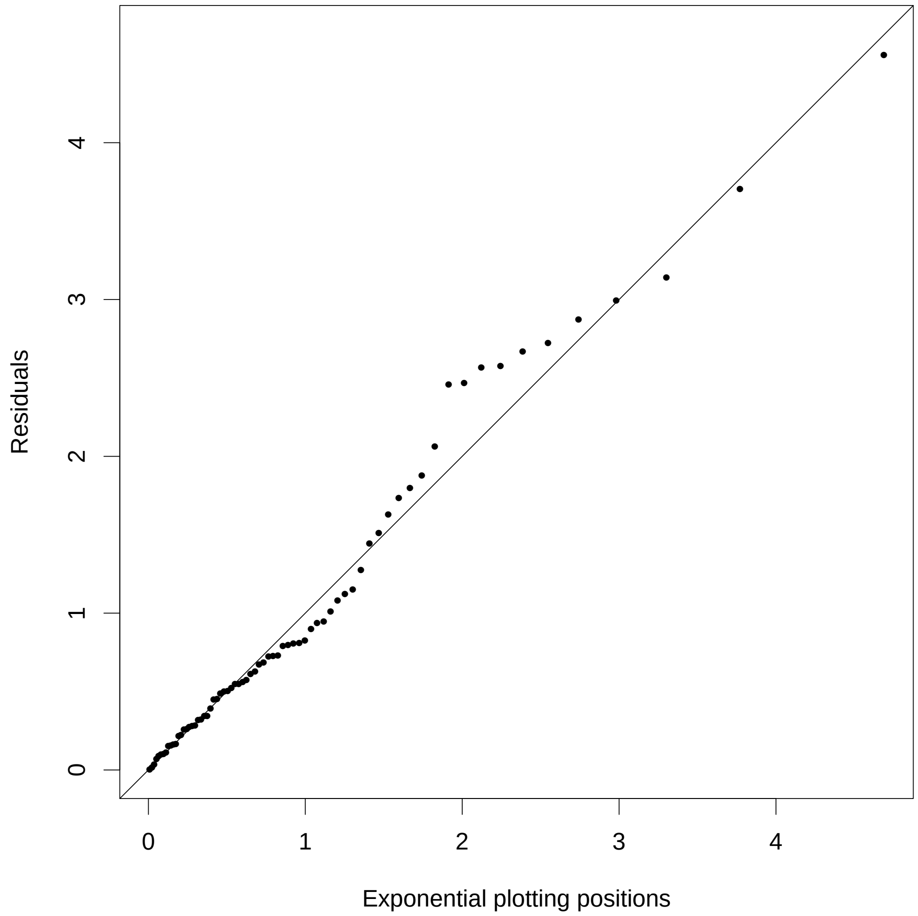


The advanced POT model: intensity function

$$\log \hat{\lambda}(t,\tau) = \hat{\gamma}_\tau I_\tau + \hat{\beta} I_{(t>t_c)} + \hat{g}(t)$$



The advanced POT model: Goodness-of-fit



The advanced POT model: (dynamic) risk measures

Some results for $\mathrm{dVaR}_\alpha^\tau(2002)$ and $\mathrm{dES}_\alpha^\tau(2002)$

	$\widehat{\mathrm{dVaR}}_{0.99}^\tau(2002)$	$\widehat{\mathrm{dES}}_{0.99}^\tau(2002)$
$\tau = 1$	138.1	556.3
$\tau = 2$	90.6	271.6
$\tau = 3$	23.9	34.9

IV. Dependence and Point Processes

Stylized situation:

$$\left.\begin{array}{ll}
\text{Loss class } 1: & L_{1,T} = \sum_{i=1}^{N_1(T)} X_{i,1} \\
\text{Loss class } 2: & L_{2,T} = \sum_{i=1}^{N_2(T)} X_{i,2} \\
 & \vdots \\
\text{Loss class } d: & L_{d,T} = \sum_{i=1}^{N_d(T)} X_{i,d}
\end{array}\right\}$$

- interdependence
- aggregation

Interdependence: available tools

- Correlation (linear, rank):
  - one-number summary: ρ, τ, ρ_S ...
- Copula: $L_{i,T} \sim F_i$
  - $\mathsf{P}(L_{1,T} \leq l_1, \ldots, L_{d,T} \leq l_d) = \mathcal{C}(F_1(l_1), \ldots, F_d(l_d))$
  - determine the joint distribution function
- Joint dynamic models for the compound processes:
  - $\{(L_{1,t}, \ldots, L_{d,t}) : t \geq 0\}$
  - marked point processes

Interdependence in the joint dynamic models

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

- 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
- via
  (1) standard copula techniques
  (2) for t fixed, copulas of discrete distributions
- however: we want to couple the dynamic processes

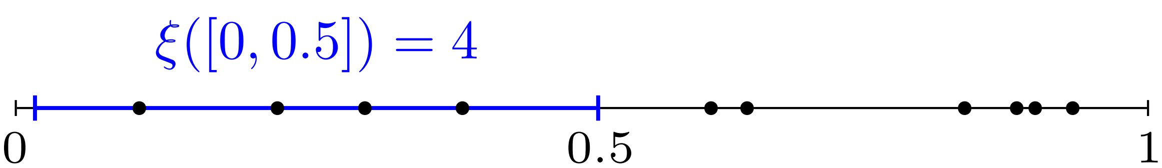
An introduction of dynamic dependence models for point processes

1. Dependent loss frequencies (time dependence)

$$\xi = \sum_{i=1}^{N} I_{\boldsymbol{T}_i}, \quad \boldsymbol{T}_i = (T_{i,1}, \ldots, T_{i,d}) \in \mathbb{R}_+^d$$

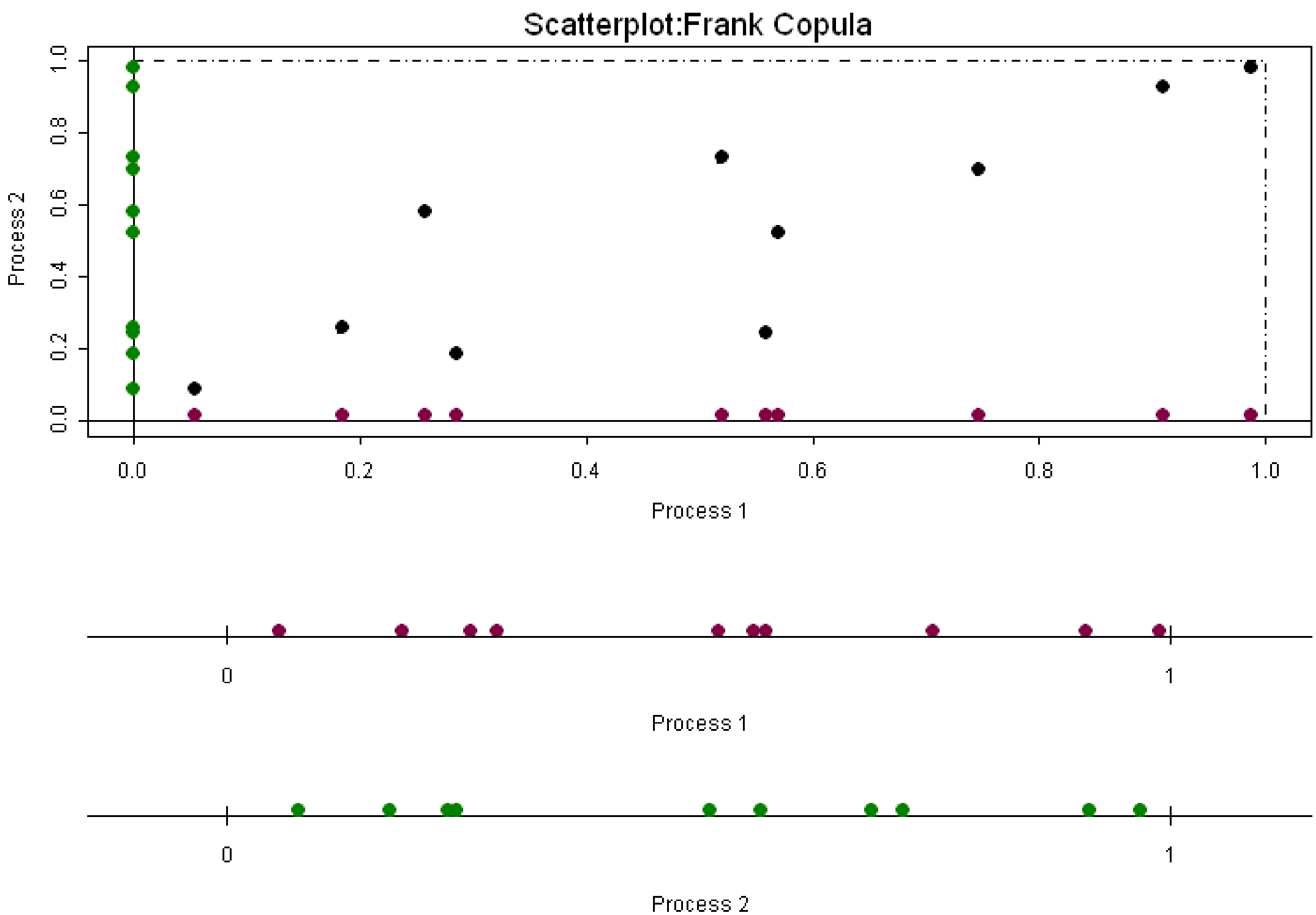
- N gives the total (random) number of occurrences
- The $\boldsymbol{T}_i$'s are the exact locations
- In the Poisson case, N is Poisson, N and $\{\boldsymbol{T}_i\}_i$ are independent and the $\boldsymbol{T}_i$'s are iid

An illustration for $d = 1$:



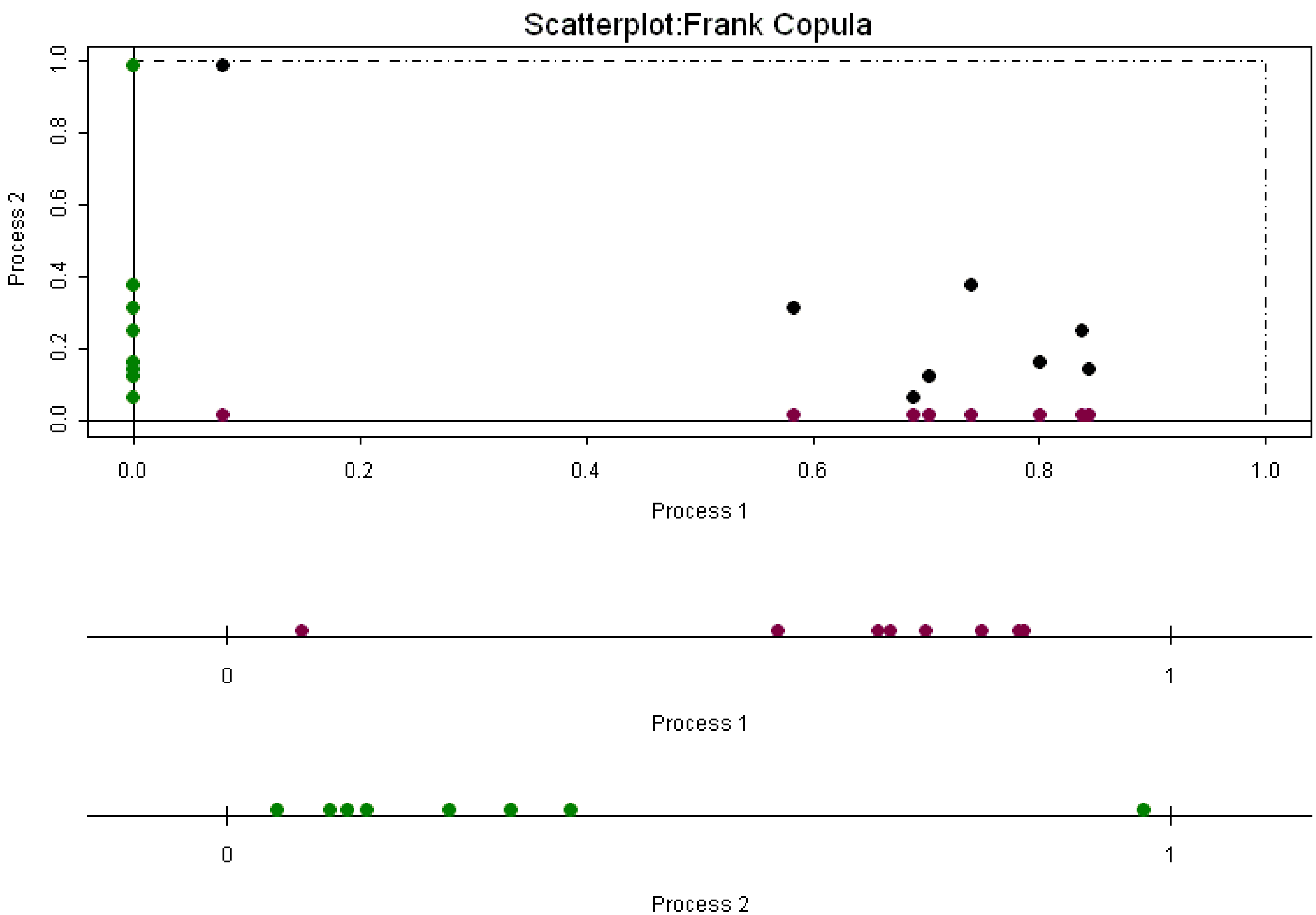
Moving to $d \geq 2$: Method I – Projections

Example 1: positive dependence (Frank copula with parameter $\theta = 20$)



Moving to $d \geq 2$: Method I – Projections

Example 2: negative dependence (Frank copula with parameter $\theta = -20$)



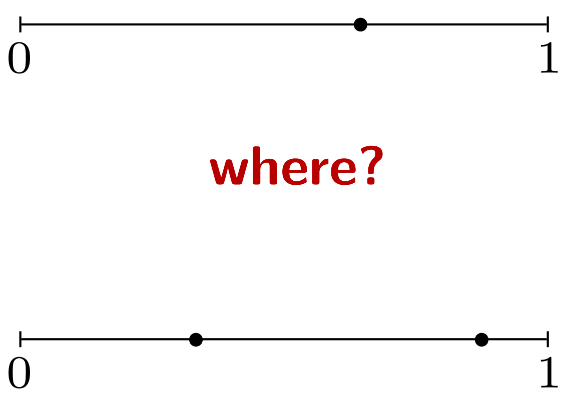
Method I: Discussion

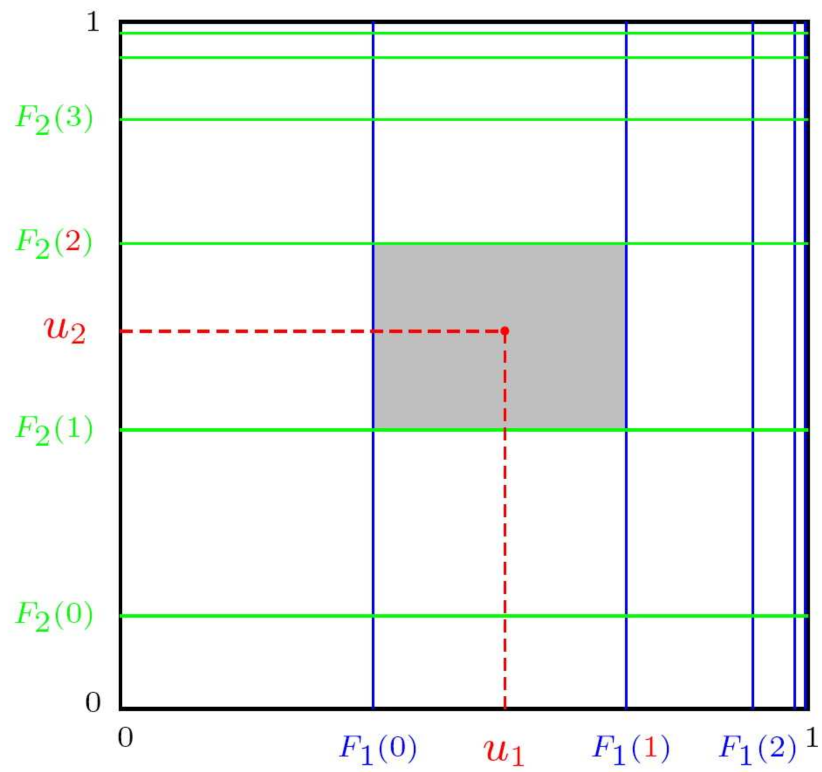
- Allows for construction of “dependence scenarios” like “dependence engineering”
- Construction holds for arbitrary d
- Does not allow for different frequencies across d classes
- Common shock scenarios
- Resulting processes are always positively dependent, the same holds for aggregated loss processes

Moving to different frequencies: Method II

- $N \quad \Rightarrow \quad (N_1, \ldots, N_d)$

- $N_1, \ldots, N_d$ dependent (copulas)

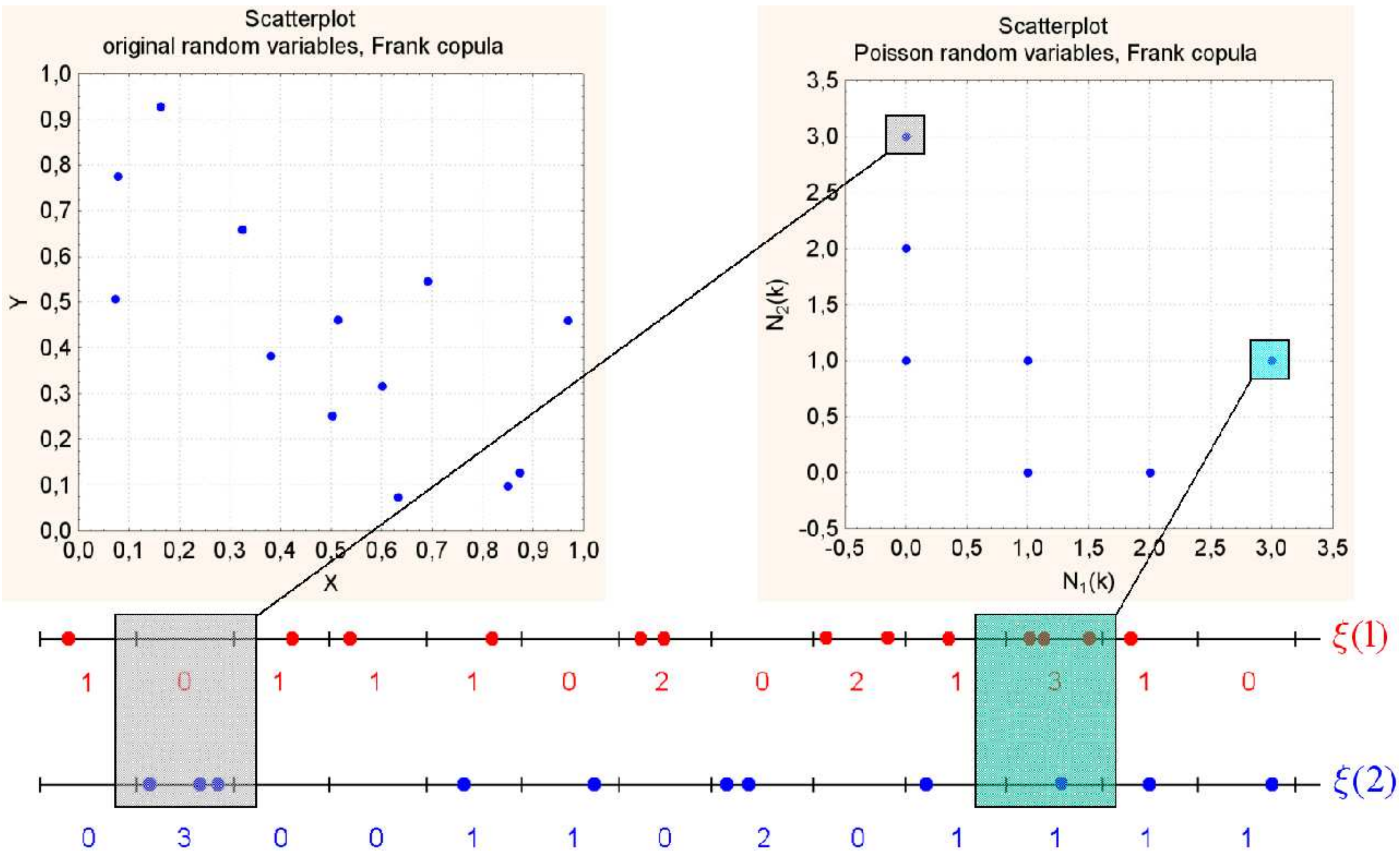




Moving to different frequencies: Method II cont'd

Be careful:

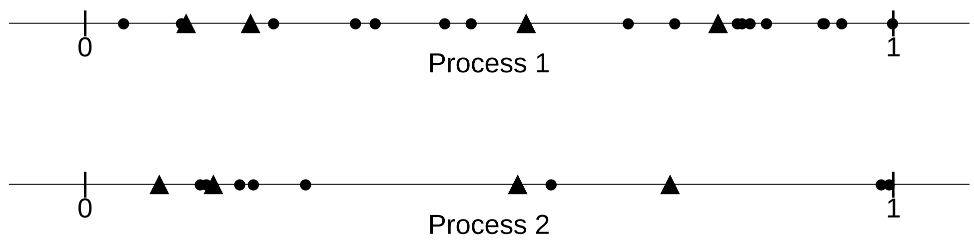
- dependence between number of events (above construction)
- locations
  - independence: standard
  - dependence: possible but tricky



Method II: Discussion

As in previous discussion, however

- Method II does allow for different frequencies across d classes
- Resulting processes can be negatively dependent, even if the locations are independent
- Method I and Method II can be combined (superposition of processes)



An introduction of dynamic dependence models for point processes

2. Including the loss-sizes

$$\left.\begin{array}{ll} \text{Loss class } 1: & L_{1,T} = \sum_{i=1}^{N_1(T)} X_{i,1} \\ \text{Loss class } 2: & L_{2,T} = \sum_{i=1}^{N_2(T)} X_{i,2} \\ & \vdots \\ \text{Loss class } d: & L_{d,T} = \sum_{i=1}^{N_d(T)} X_{i,d} \end{array}\right\}$$

- within independent
- across dependent

- "dependence across" comes from the interdependence of the corresponding loss frequencies
  - Method I
  - Method II
  - Combination of both

Dependent loss processes

- Method I ➡ $\rho(L_{1,T}, L_{2,T}) = \frac{\mathrm{E}(X_{1,1}X_{1,2})}{\sqrt{\mathrm{E}(X_{1,1})^2\,\mathrm{E}(X_{1,2})^2}}$

- Method II ➡ $\rho(L_{1,T}, L_{2,T}) = \rho(N_1(T), N_2(T))\frac{\mathrm{E}(X_{1,1}X_{1,2})}{\sqrt{\mathrm{E}(X_{1,1})^2\,\mathrm{E}(X_{1,2})^2}}$

Frachot, A., Roncalli, T. and Salomon, E. (2004) *The correlation problem in operational risk.* Crédit Lyonnais, Working paper.

- Combination of Method I and Method II ➡ common shock model as a special case

Powojowski, M., Reynolds, D. and Tuenter, H. (2002) Dependent events and operational risk. *Algo research quarterly*, **5**(2), 68-73

V. Aggregation

- In practice: $d \in \{7, 8, 56\}$
- Loss rvs: $L_{1,T}, \dots, L_{d,T}$ dependent
- With given risk measures: $\mathrm{VaR}^{\mathrm{T}}_{1,\alpha}, \dots, \mathrm{VaR}^{\mathrm{T}}_{\mathrm{d},\alpha}$
- Issue:

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

Under some (or no) idea of the interdependence

Question 1

$$\mathrm{VaR}_{\alpha}^{\mathrm{T}}\left(\sum_{\mathrm{k}=1}^{\mathrm{d}} \mathrm{L}_{\mathrm{k},\mathrm{T}}\right) \le \sum_{\mathrm{k}=1}^{\mathrm{d}} \mathrm{VaR}_{\mathrm{k},\alpha}^{\mathrm{T}} \quad ?$$

No in general:

- highly skewed loss dfs
- (very) heavy-tailed loss dfs
- special dependence structures

This is an issue in OpRisk!

Question 2

Find optimal bounds for

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

given marginal VaR's and dependence information

Solution:

- Fréchet Problem
- Mass Transportation Problem

Example 1: Comonotonic risks

If $L_{1,T}, \ldots, L_{d,T}$ are comonotonic, i.e. there exists a rv Z and increasing functions $f_{1,T}, \ldots, f_{d,T}$, so that

$$L_{i,T} = f_{i,T}(Z) \qquad i = 1, \ldots, d,$$

then the VaR is additive, i.e.

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

Example 2: No dependence information

(Unconstrained optimization problem)

- Take $L_{i,T} = L_i, \quad i = 1, \ldots, d = 8,$
  - Marginal OpRisk loss dfs:

$$P(L_i \le x) = 1 - (x+1)^{-1.5}, \quad x \ge 0 \quad (\mathrm{Pareto}(1, 1.5))$$

  - $E(L_i) = 2 < \infty, \quad \mathrm{Var}(L_i) = \infty$

- In the comonotonic case:

$$\mathrm{VaR}_{0.99}\left(\sum_{i=1}^{8} L_i\right) = \sum_{i=1}^{8} \mathrm{VaR}_{0.99}(L_i) = 0.16$$

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

Example 2: No dependence information cont'd

- The unconstrained bounds are (in thousand):

$$\mathrm{VaR}_{0.99}\left(\sum_{i=1}^{8} L_i\right) \leq 0.41$$

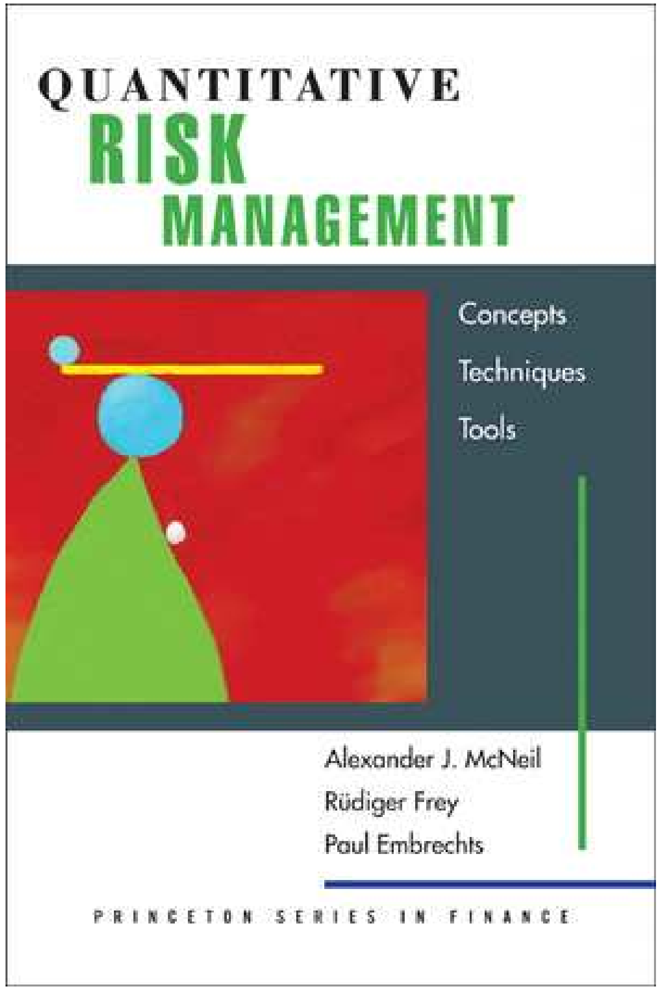
$$\mathrm{VaR}_{0.999}\left(\sum_{i=1}^{8} L_i\right) \leq 1.93$$

- The "worst dependence case" can be calculated

VI. Conclusion

- OpRisk data are intricate
- Regulators ask for extreme risk measures
- At the moment we see no satisfactory full AMA model at the horizon
- There are interesting methodological building blocks working well for specific OpRisk data
- There are interesting more advanced techniques
- More work is needed
- DATA!

Quantitative Risk Management: Concepts, Techniques and Tools



References

Chavez-Demoulin, V. and Embrechts, P. (2004) Smooth Extremal Models in Finance. *The Journal of Risk and Insurance* **71**(2), 183–199.

Chavez-Demoulin, V. and Davison, A. C. (2005) Generalized additive models for sample extremes. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **54**(1), 207–222.

Embrechts, P. and Puccetti, G. (2005) Bounds for functions of dependent risks. *Finance and Stochastics*, to appear.

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

Neslehova, J. and Pfeifer, D. (2004) Modeling and generating risk processes for IRM and DFA. *Astin Bulletin*, **34**(2), 333–360.