Quantitative Modeling of Operational Risk: Between g-and-h and EVT

Paul Embrechts Matthias Degen Dominik Lambrigger

ETH Zurich
(www.math.ethz.ch/∼embrechts)

Outline

What is Basel II?

- **1988 Basel I Accord on Banking Supervision**
  - mainly CR
  - minimum risk capital (MRC) $\geq$ 8% of risk weighted assets (Cooke Ratio)

- **1993 Birth of VaR**
  - "G-30 Report" addressing incorporation of off-balance sheet products (first time "VaR" appears)
  - need for proper RM of these products

- **1996 Amendment to Basel I**
  - standardized model for MR
  - internal models allowed
  - legal implementation in 2000

- **2001 Initiation of consultative process for Basel II**
  - refined CR-approaches, IRB-models
  - consideration of new risk class: OR
  - implementation 2007+

▶ note Solvency I & II

Risk Components (Basel II)

- **Credit Risk**
- **Market Risk**
- **Operational Risk**
- **Business Risk**

Risk Components (Basel II)

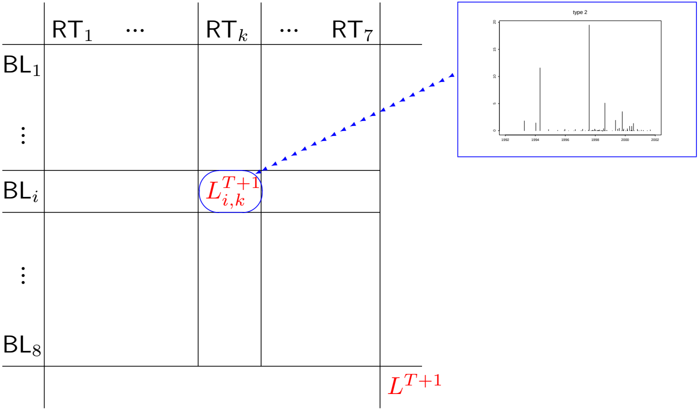
- **Credit Risk**
- **Market Risk**
- **Operational Risk**
- **Business Risk**

Operational Risk: The risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. Including legal risk, but excluding strategic and reputational risk.

Some examples

- 1995: Nick Leeson/Barings Bank, £1.3b
- 2001: September 11
- 2001: Enron (largest US bankruptcy so far)
- "Fat finger" errors

Loss Distribution Approach (LDA)



Basel II - Guidelines

- **Risk measure:** VaR
- **Time horizon:** 1 year
- **Level:** 99.9% (1 in 1000 year event!)

▶ **Otherwise:** Full methodological freedom (within LDA)

The Main LDA-Steps towards a Total Capital Charge

- Estimation of marginal VaR: $$\widehat{\mathrm{VaR}}_\alpha^1, \ldots, \widehat{\mathrm{VaR}}_\alpha^d$$
- Additional Aggregation: $$\widehat{\mathrm{VaR}}_\alpha^+ = \sum_{k=1}^{d} \widehat{\mathrm{VaR}}_\alpha^k$$
- Diversification: $$\mathrm{VaR}_\alpha^{\mathrm{real}} \overset{?}{<} \widehat{\mathrm{VaR}}_\alpha^+$$

Reasonable Severity Distribution*

- **Good statistical fit of the data**
- **Loss distribution with realistic capital estimates**
- **Well specified:** Are the characteristics of the fitted data similar to the loss data and logically consistent?
- **Flexible:** How well is the method able to reasonably accomodate a wide variety of empirical loss data?
- **Simple:** Is the method easy to apply in practice?

*see Dutta and Perry (2006)

Loss Distribution

EVT	**g-and-h**
• Moscadelli (2004): – reasonable capital estimates (LDCE 2002) – infinite mean models occur	• Dutta and Perry (2006): – EVT fails, propose g-and-h (LDCE 2004) – finite mean g-and-h models
• Well established theory: Peaks Over Threshold (POT)	• No standard framework (yet)
• No specific underlying df	• Specific parametric model

▶ Careful look at the g-and-h approach

g-and-h: Basic Properties

Definition

Let $Z \sim \mathcal{N}(0,1)$ be a standard normal rv. A rv X is said to have a g-and-h distribution with parameters $a, b, g, h \in \mathbb{R}$, if X satisfies

$$\boxed{X = k(Z) = a + b\frac{e^{gZ} - 1}{g} e^{hZ^2/2}}$$

- g governs skewness
- h governs heavy-tailedness
- Distributional properties of $F \sim$ g-and-h?

Theorem 1

Suppose $F \sim$ g-and-h, then:

- For $g, h > 0$, we have $\overline{F} \in RV_{-1/h}$, i.e. $\overline{F}(x) = x^{-1/h} L(x)$ with $L \in SV$.
- For $h = 0$ and $g > 0$, we have $F \in \mathcal{S} \backslash RV$, where $\mathcal{S}$ denotes the class of subexponential dfs.

▶ Well-known theory of (1st and 2nd order!) regular variation

Theorem 2

The slowly varying function L asymptotically behaves like

$$\frac{\exp\left(\sqrt{\log x}\right)}{\sqrt{\log x}}, \quad x \to \infty.$$

- Difficult type of slowly varying function

Pickands-Balkema-de Haan Theorem

First order property:

$$\lim_{u \uparrow x_0} \underbrace{\sup_{x \in (0, x_0 - u)} \left| F_u(x) - G_{\xi, \beta(u)}(x) \right|}_{=: d(u)} = 0$$

- $F_u(x) = P(X - u \leq x | X > u)$: excess df
- $G_{\xi, \beta(u)}$: generalized Pareto distribution (GPD)
- $x_0 \leq \infty$: upper endpoint

Pickands-Balkema-de Haan Theorem (continued)

- Theory: Under weak conditions $d(u)$ converges to 0.
  (Maximum Domain of Attraction)
- Practice: No information on goodness of approximation.

Pickands-Balkema-de Haan Theorem (continued)

- Theory: Under weak conditions $d(u)$ converges to 0.
  (Maximum Domain of Attraction)

- Practice: No information on goodness of approximation.

Second order property:

- How fast does $d(u)$ converge to 0?
- Determined by $L \in SV$
- Highly relevant for practical applications

Rate of convergence to the GPD for different distributions, as a function of the threshold u

Distribution	Parameters	$\overline{F}$	$d(u)$
Exponential(λ)	$\lambda > 0$	$e^{-\lambda x}$	0
Pareto(α)	$\alpha > 0$	$x^{-\alpha}$	0
Double exp. parent		$e^{-e^{x}}$	$O(e^{-u})$
Student t	$\nu > 0$	$\overline{t}_{\nu}(x)$	$O(\frac{1}{u^2})$
Normal(0, 1)		$\overline{\Phi}(x)$	$O(\frac{1}{u^2})$
Weibull(τ, c)	$\tau \in \mathbb{R}_{+} \setminus \{1\}, c > 0$	$e^{-(cx)^{\tau}}$	$O(\frac{1}{u^{\tau}})$
Lognormal(μ, σ)	$\mu \in \mathbb{R}, \sigma > 0$	$\overline{\Phi}(\frac{\log x - \mu}{\sigma})$	$O(\frac{1}{\log u})$
Loggamma(γ, α)	$\alpha > 0, \gamma \neq 1$	$\overline{\Gamma}_{\alpha,\gamma}(x)$	$O(\frac{1}{\log u})$

Rate of convergence to the GPD for different distributions, as a function of the threshold u

Distribution	Parameters	$\overline{F}$	$d(u)$
Exponential(λ)	$\lambda > 0$	$e^{-\lambda x}$	0
Pareto(α)	$\alpha > 0$	$x^{-\alpha}$	0
Double exp. parent		$e^{-e^{x}}$	$O(e^{-u})$
Student t	$\nu > 0$	$\overline{t}_{\nu}(x)$	$O(\frac{1}{u^2})$
Normal$(0,1)$		$\overline{\Phi}(x)$	$O(\frac{1}{u^2})$
Weibull(τ, c)	$\tau \in \mathbb{R}_+ \setminus \{1\}, c > 0$	$e^{-(cx)^{\tau}}$	$O(\frac{1}{u^{\tau}})$
Lognormal(μ, σ)	$\mu \in \mathbb{R}, \sigma > 0$	$\overline{\Phi}(\frac{\log x - \mu}{\sigma})$	$O(\frac{1}{\log u})$
Loggamma(γ, α)	$\alpha > 0, \gamma \neq 1$	$\overline{\Gamma}_{\alpha,\gamma}(x)$	$O(\frac{1}{\log u})$
g-and-h	$g, h > 0$	$\overline{\Phi}(k^{-1}(x))$	$O(\frac{1}{\sqrt{\log u}})$

Rate of convergence to the GPD for different distributions, as a function of the threshold u

Distribution	Parameters	$\overline{F}$	$d(u)$
Exponential(λ)	$\lambda > 0$	$e^{-\lambda x}$	0
Pareto(α)	$\alpha > 0$	$x^{-\alpha}$	0
Double exp. parent		e^{-e^x}	$O(e^{-u})$
Student t	$\nu > 0$	$\overline{t}_\nu(x)$	$O(\frac{1}{u^2})$
Normal$(0,1)$		$\overline{\Phi}(x)$	$O(\frac{1}{u^2})$
Weibull(τ, c)	$\tau \in \mathbb{R}_+ \setminus \{1\}, c > 0$	$e^{-(cx)^\tau}$	$O(\frac{1}{u^\tau})$
Lognormal(μ, σ)	$\mu \in \mathbb{R}, \sigma > 0$	$\overline{\Phi}(\frac{\log x - \mu}{\sigma})$	$O(\frac{1}{\log u})$
Loggamma(γ, α)	$\alpha > 0, \gamma \neq 1$	$\overline{\Gamma}_{\alpha,\gamma}(x)$	$O(\frac{1}{\log u})$
g-and-h	$g, h > 0$	$\overline{\Phi}(k^{-1}(x))$	$O(\frac{1}{\sqrt{\log u}})$

- If data are well modeled by a g-and-h, EVT-based estimation converges very slowly

Tail Index Estimation

- $X_i \overset{iid}{\sim} \overline{F} \in RV_{-1/\xi}$
- $H_{k,n} := \frac{1}{k} \sum_{j=1}^{k} (\log X_{n-j+1,n} - \log X_{n-k,n})$ (Hill estimator)
- $H_{k,n}$ very sensitive to choice of threshold k
- "optimal" k often s.t. AMSE of $H_{k,n}$ minimal

Tail Index Estimation - Simulation Study

heavy-tailedness →

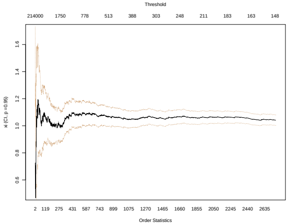
skewness ↓

$g \backslash h$	0.1	0.2	0.5	0.7	1	2
0.1	142	82	33	23	18	11
0.2	165	97	42	32	25	20
0.5	224	132	49	38	27	19
0.7	307	170	63	44	29	20
1	369	218	86	58	36	26
2	696	385	151	108	74	31
3	1097	613	243	163	115	54

Empirical SRMSE (in %) of the Hill estimator $\widehat{h}^{Hill}_{kopt}$ of h for g-and-h data for different parameter values of g and h

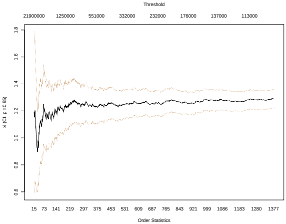
Hill Plots

$\frac{g}{h}$ small



- Hill plot works fine $(g = 0.1, h = 1)$

$\frac{g}{h}$ large



- Hill plot misleadingly indicates infinite mean model! $(g = 4, h = 0.2)$

Aggregation

Dutta-Perry: "We have not mathematically verified the subadditivity property for g-and-h, but in all cases we have observed empirically that enterprise level capital is less than or equal to the sum of the capitals from business lines or event types."

Question: $$C_\alpha^{\text{OpRisk}} < \widehat{\text{VaR}}_\alpha^{+} \stackrel{\text{def}}{=} \sum_{k=1}^{d} \widehat{\text{VaR}}_\alpha^{k} \quad ?$$

Subadditivity of VaR typically fails for:

- Skewness
- Heavy-Tailedness
- Dependence

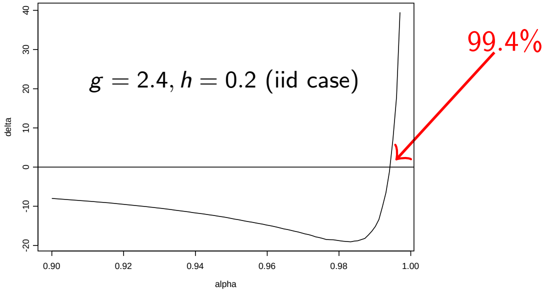
Remark

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

- ▶ No reasonable risk measures exist
- ▶ Diversification goes the wrong way

Proposition [Daníelsson et al.]

Suppose that the non-degenerate vector (X_1, X_2) is regularly varying with extreme value index $\xi < 1$. Then VaR_α is subadditive for α sufficiently large.



diversification benefit:

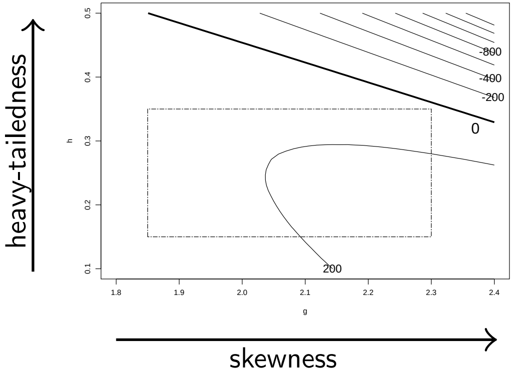
$$\text{delta} = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2) - \mathrm{VaR}_\alpha(X_1 + X_2)$$

Remark

This proposition is only an asymptotic statement - It does not guarantee subadditivity for a broad range of high quantiles

- of no use for practical assessment of subadditivity
- Basel II: 1-year 99.9% VaR - which choices of g and h yield subadditive models?

Subadditivity of VaR at 99.9%

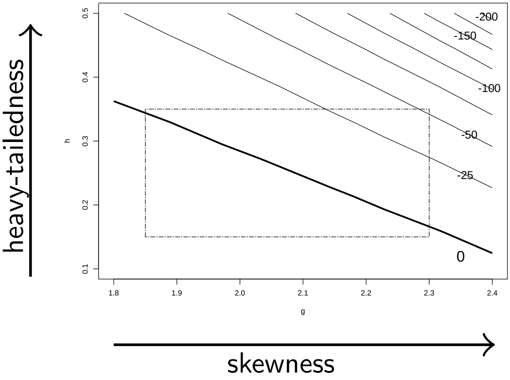


- Entire parameter rectangle within subadditivity range
- Small changes of parameters ⇒ superadditivity

What happens when we go deeper in the data?

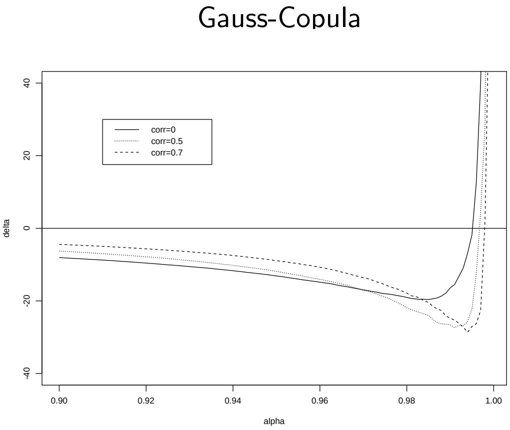
- VaR-estimation at 99.9% and higher: difficult!
- Estimate at lower level (90%, say) and scale: how?

Subadditivity of VaR at 99%



- Substantial fraction of parameter rectangle switched regime
- Far from diversification!

Dependence matters



Increasing correlation ⇒ superadditivity range extends

Conclusion

- Very slow convergence of g-and-h excess df to the GPD when $g, h > 0$
- Optimal threshold selection for an EVT based POT approach becomes very difficult (unreliable risk capital estimates)
- Small changes of g and/or h may lead VaR to switch (sub-/superadditivity) regime
- g-and-h is subexponential $\longrightarrow$ one claim causes ruin

References

- Degen, M., Embrechts, P. and Lambrigger, D. (2006) The quantitative modeling of operational risk: between g-and-h and EVT. ASTIN Bulletin 2007, to appear.
- Dutta, K. and Perry, J. (2006) A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital. Federal Reserve Bank of Boston, Working Paper No 06-13.
- Moscadelli, M. (2004) The modelling of operational risk: experiences with the analysis of the data collected by the Basel Committee. Bank of Italy, Working Paper No 517.