IAA ASTIN Group Webinar, 21 Oct. 2021

Three-layer problems and the Generalized Pareto distribution

Michael Fackler

Independent Consulting Actuary,

Munich, Germany

About the speaker

- Dr. (rer. nat.) Michael Fackler
- Qualified actuary (DAV), self-employed
- Studied Math at Univ. Munich, Pisa, Oldenburg
- Doctorate in parallel with working: on experience rating, completed 2017
- 10 years with leading reinsurers
- 15+ years as consulting actuary
- Specialized in: non-life reinsurance pricing, dealing with scarce data

Outline

- Introduction
- Generalized Pareto
- 3-layer problem
- Variants
- Model risk

Situation

Tail modelling, e.g. for layer pricing, Solvency

- Very scarce loss data
- Helpful information possibly from different sources, e.g. *your portfolio* vs *market benchmark*
- Models not fully specified
- Only easily accessible data bits:

frequencies at thresholds / risk premiums of layers

Examples: 3 reinsurance layers

Liability / Fire per risk, NatCat, etc.

- layer 1: experience rate
- layer 2: blend experience / exposure
- layer 3: exposure (first loss curves / vendor model)

MTPL experience rates

- layer 1: company
- layer 2: country
- layer 3: continent

MTPL example

Task: Pricing of layers from **1** up to **20** (mln USD)

Input:

- A dozen large losses from your portfolio enable you to quote the layer **2 xs 1**, risk premium: **1.04**
- For the whole market someone quoted the layer **5 xs 5**, risk premium: **3**
- Your portfolio supposedly has average exposure, market share is 8%, thus your risk premium for this ”market” layer would be: **0.24**

MTPL example

For higher layers you don’t have market quotations or don’t believe them.

Workaround:

- Maximum desired *payback* period for large events (possibly politically set): **200 years**

General approach

- Be **modest**: no *best*-fit ambitions, a *good-enough* model is fine (*satisficing*)
- Make as few assumptions as possible
- Rather assume *explicitly* a **return period** than *implicitly* a specific parametric **model** – this may not always be the better decision, but it is transparent and well testable.

Methodology

- Use the Collective Model
- Try to find frequency / severity that reproduce given data bits (moment-matching variant)
- Use the GPD (tail) severity above appropriate threshold *s*

$$P(Z > x | Z > s) = \left(\left(1 + \xi \frac{x - s}{\sigma} \right)^{+} \right)^{\frac{-1}{\xi}}$$

Why the GPD?

- Extreme Value Theory (everyone accepts GPD)
- empirical: fits well lots of data
- practical: easy to use, analytical formulae
- intuitive: Parameters can be interpreted
- tail-invariant:
  If a distribution tail is GPD, so are higher tails.

Tail-invariant parameters

- generally: $\xi, \ \sigma^* = \sigma - \xi s > -\xi s$
- $\xi > 0$: $\alpha = 1/\xi, \ \lambda = \alpha\sigma^*$; Pareto: $\lambda = 0 = \sigma^*$
- $\xi < 0$: $\beta = -\alpha, \ \nu = -\lambda$; ν is supremum loss
- $\xi = 0$: $\sigma = \sigma^*$; Exponential

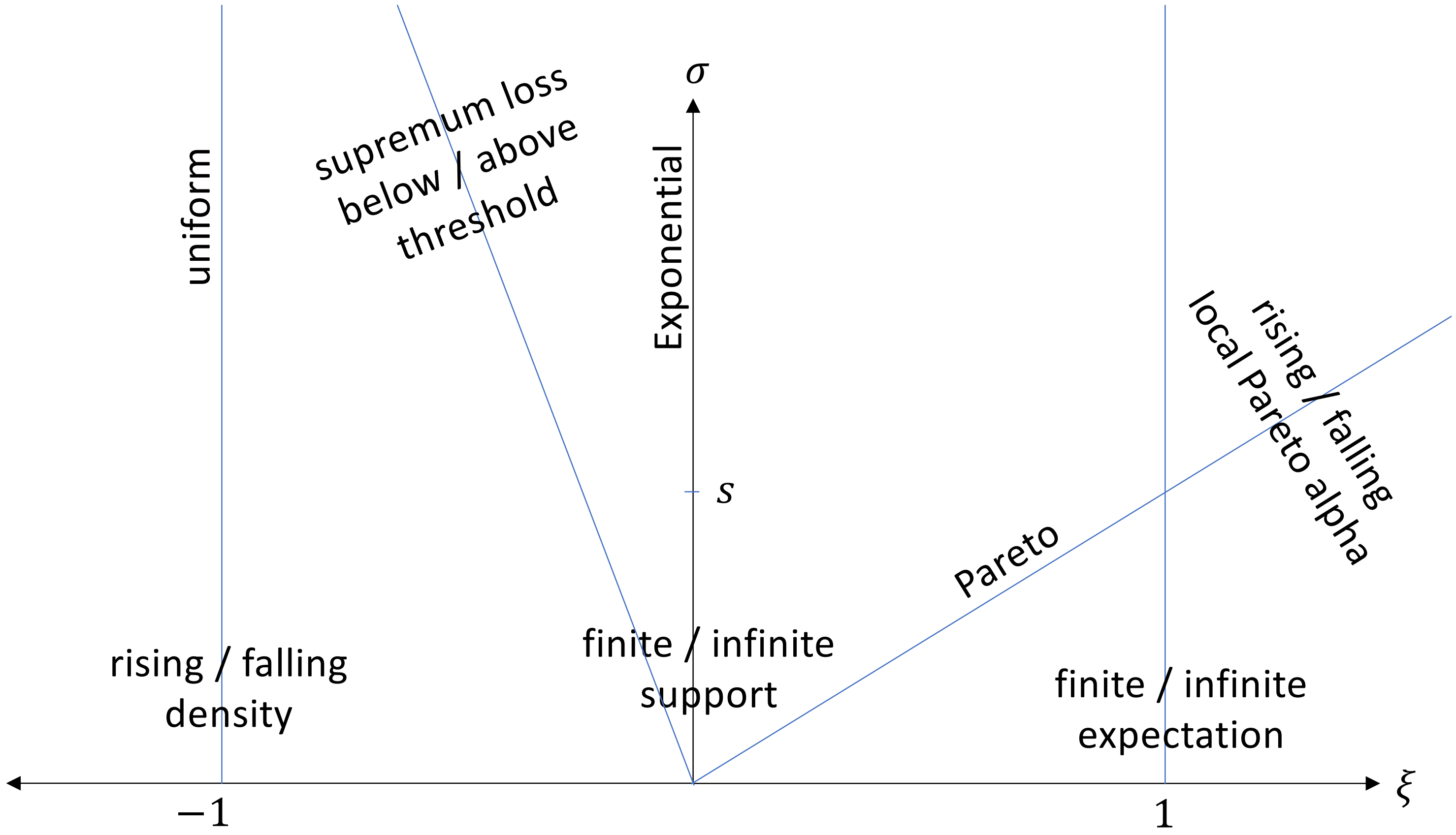
$$P(Z > x|Z > s) = \left(\frac{\xi s + \sigma^*}{(\xi x + \sigma^*)^+}\right)^{1/\xi} =$$

$$\left(\frac{s+\lambda}{x+\lambda}\right)^{\alpha} \text{ vs. } \left(\frac{(\nu - x)^+}{\nu - s}\right)^{\beta} \text{ vs. } \exp\left(\frac{x-s}{\sigma}\right)$$

Local Pareto alpha

- Approximates cdf $F(x)$ locally by Pareto curve
- General definition: $\alpha_d := d\frac{f(d)}{1-F(d)}$
- for the GPD $\alpha_d = \frac{d}{\sigma+\xi(d-s)}$
- for $\xi > 0$ $\alpha_d = \alpha\frac{d}{d+\lambda}$

GPD map



uniform
supremum loss
below / above
threshold
σ
Exponential
s
rising / falling
local Pareto alpha
Pareto
rising / falling
density
finite / infinite
support
finite / infinite
expectation
-1
1
ξ

Three-layer problems

Given input:

- Risk premiums for 3 “reference” layers
- Frequencies for 3 thresholds
- Mixed cases

Threshold is limiting case of layers:
frequency at threshold = *risk rate on line* of very thin layer

$$RRoL = \frac{risk\ premium}{layer\ limit}$$

MTPL example

Formulate as (mixed) three-layer problem:

- layer **2 xs 1**: *RRoL* = **52%**
- layer **5 xs 5**: *RRoL* = **4.8%**
- threshold **20**: *freq.* = **0.5%**

Theorem

For 3 non-overlapping layers with given RRoL's

$$r_1 > r_2 > r_3 > 0$$

the 3-layer problem can (mostly) be solved:

by a **unique** GPD severity

together with a (unique) frequency f at the attachment point s of the lowest layer

Theorem

- Works also with thresholds or mixed input
- $s = 0$ is possible (ground-up model)
- Top layer may be unlimited ($r_3 = 0$)
- Layers 1 and 2 may overlap (to some extent)
- Uniqueness holds for further situations

Case with no solution:

- Layer 1 is no threshold, Layer 3 is limited
- $r_1 \gg r_2 \approx r_3$

MTPL example

$s = 1$ (million USD)

- $f = 1.09$
- $\xi = 0.41$ $\qquad$ $(\alpha = 2.44)$
- $\sigma = 0.96$

Remarks

- Model is easy to find numerically:
  search 3 parameters fulfilling 3 RRoL equations
- Special case *layer-endpoints problem*:
  one layer with 3 data bits:
  risk premium, layer entry / exit frequencies
- Single-parameter Pareto solves analogous
  2-layer problems
- GPD solves many real-world **4-layer problems**
  approximately

Remarks

Paper gives **model-building recipes** for a variety of scarce-data situations

- Piecewise GPD solves **n-layer problems** exactly

Procedure:

- GPD for highest 3 (or less) layers
- Attach layers downwards via layer-endpoint problem

MTPL example variant

- Layer 1 **2 xs 1**: *RRoL* = **52%**
- Layer 2 **2 xs 3**: *RRoL* = **13%**
- Layer 3 **5 xs 5**: *RRoL* = **4.8%**
- Layer 4 **10 xs 10**: *RRoL* = **1.3%**

Approximate GPD ”fit”:

- 3 parameters, 4 equations, and a deviation metric, e.g. sum of squared RRoL deviations (absolute, relative, or something intermediate)

MTPL example variant

Exact piecewise-GPD ”fit”:

- GPD for Layers 2 - 4
- Resulting loss frequency at 3 is 19%, which is consistent with RRoL 52% of Layer 1
- Choose consistent frequency at 1 (>52%) and solve layer-endpoint problem for Layer 1

Model risk

… must be high with scarce data, however:

- **Major uncertainty** is expected loss – and possibly the loss count model
- Higher moments of the severity often don’t bear much further uncertainty, in particular for layers in the middle of a program
- GPD is a *choice*, but a good one, both in **practical** and **statistical** sense: other severities are less handy and will often produce very similar output

Parameter-free inequality

Limited layer: limit c, layer severity X,

$$f \geq r \geq g \geq 0$$

with loss frequency f, total loss freq. g, RRoL r:

$$1 - \frac{f-r}{f-g}\frac{r-g}{r} \leq \frac{\mathrm{E}(X^2)}{c\,\mathrm{E}(X)} \leq 1$$

- Interval is narrow for heavy-tailed severity
- Narrower interval for concave cdf
- Analogous bounds for higher moments

Parameter-free inequality

Respective interval for the aggregate layer loss S follows from

$$\mathrm{CV}^2(S) = \mathrm{Ct}(N) + \frac{1}{r}\,\frac{\mathrm{E}(X^2)}{c\,\mathrm{E}(X)}$$

with the *contagion* of the loss count

$$\mathrm{Ct}(N) = \mathrm{CV}^2(N) - \frac{1}{\mathrm{E}(N)} = \frac{1}{\mathrm{E}(N)}\left(\frac{\mathrm{Var}(N)}{\mathrm{E}(N)} - 1\right)$$

a most useful quantity

Wrap up

If you want to extrapolate

- below the lowest or
- above the highest

of your reference layers, results inevitably depend strongly on the choice of the parametric model.

Instead, for all layers somewhere between the reference layers model risk is very low.

Conclusion

The building of models by solving three-layer problems is **powerful** and, in case of very scarce data, an excellent **trade-off** between *statistical ambition* and the *need to get things done*.

Thanks for joining this talk. See the paper for details (2021 ASTIN Colloquium website or SSRN).

michael_fackler@web.de