DAV-Herbsttagung, ASTIN-Gruppe
Mainz / Online, 14.11.2022

Das Drei-Layer-Problem

oder

Three-layer problems and the Generalized Pareto distribution

(ASTIN Colloquium 2021, Best Paper Award)

Dr. Michael Fackler

freier Aktuar, München

DAV/DGVFM
Herbsttagung
2022

Abstract

Der klassische Weg, eine parametrische Verteilung für Versicherungs-Großschäden zu finden, ist die Parameterschätzung aus empirischen Schadendaten. In der Praxis reichen die Daten hierfür oft nicht aus, aber trotzdem wird ein parametrisches Modell benötigt, z.B. für die Rückversicherung oder Solvenz-Fragen.

Der Vortrag stellt eine Methode vor, wie man ein Generalized-Pareto-Modell aus ganz wenigen Daten konstruieren kann: den Frequenzen von 3 Schadenhöhen, den Risikoprämien von 3 Layern, oder einer Mischung solcher Inputs.

Es zeigt sich, dass passende GPD-„Fits“ bei realistischem Dateninput existieren und eindeutig sind. Dazu werden keine weiteren Informationen benötigt und man kann ein konsistentes Modell auch dann konstruieren, wenn die verwendeten 3 Inputs aus verschiedenen Quellen stammen, z.B. teils aus individuellen, teils aus Marktdaten.

Die Methode ist leicht zu implementieren und auf Situationen mit etwas weniger oder mehr als 3 Dateninputs erweiterbar.

Gliederung

- Einführung
- Generalized Pareto
- Drei-Layer-Problem
- Varianten
- Modellrisiko

Situation

Tail-Modellierung benötigt, z.B. für Layer-Prämien oder Risikokapital-Modellierung

- Sehr wenig Schadendaten
- Nützliche Info aus verschiedenen Quellen, z.B. *eigenes Portefeuille bzw. Marktstatistik*
- Modelle evtl. nicht komplett spezifiziert
- Einzige leicht zugängliche Daten-Inputs:

  Großschaden-Frequenzen / Layer-Risikoprämien

Mögliche Inputs für 3 Rückversicherungs-Layerprämien

KH Burning-Cost-Raten

- Layer 1: Versicherer
- Layer 2: Land
- Layer 3: Kontinent

Haftpflicht / Feuer pro Risiko, NatCat, etc.

- Layer 1: Experience (Burning Cost / Fit)
- Layer 2: Mischung Experience / Exposure
- Layer 3: Exposure (-kurven / geophysik. Modell)

Beispiel KH

Ziel: Preise für Layer von **1** bis **20** (Mio. USD)

Input:

- Ein Dutzend Großschäden des eigenen Pf. ermöglichen Kalkulation für den Layer **2 xs 1**, Risikoprämie: **1,04**
- Für den Gesamtmarkt gibt es eine Quotierung des Layers **5 xs 5**, Risikoprämie: **3**
- Das eigene Pf. ist markttypisch, hat Marktanteil 8%. Das ergibt für den „Markt”-Layer die Risikoprämie: **0,24**

Beispiel KH

Für höhere Layer gibt es keine Marktquotierungen oder man glaubt ihnen nicht.

Notlösung:

- Maximaler akzeptierter *Payback* für extreme Schäden (evtl. politisch entschieden): **200 Jahre**
- Wird angesetzt für bestimmte Schadenhöhe, hier z.B. für **20**

Allgemeiner Ansatz

- **Bescheidenheit**: keine *Best*-Fit-Ambitionen, das Modell muss *gut genug* sein („*Satisficing*“)
- So wenige Annahmen wie möglich
- Lieber *explizit* eine **Wiederkehrperiode** annehmen als *implizit* ein spezielles parametrisches **Modell**
  – das ist sicher nicht immer die bessere Entscheidung, aber es ist transparent und testbar.

Methode

- Kollektives Modell
- Suche Schadenfrequenz und -höhe, die zum gegebenen Dateninput passen
(im Grunde eine Momenten-Methode)
- Modelliere Schadenhöhen(-Tail) mit der GPD ab einem geeignetem Threshold *s*

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

Warum die GPD?

- Extremwerttheorie (GPD allgemein akzeptiert)
- empirisch: passt zu vielen Taildaten
- praktisch: einfach anwendbar, analytische Formeln
- intuitiv: Parameter sind interpretierbar
- tailinvariant: Wenn ein Tail GPD-verteilt ist, dann auch alle höheren Tails.

Tail-invariante Parameter

- allgemein: $\xi,\ \sigma^* = \sigma - \xi s > -\xi s$
- $\xi > 0$: $\alpha = 1/\xi,\ \lambda = \alpha\sigma^*$; Pareto: $\lambda = 0 = \sigma^*$
- $\xi < 0$: $\beta = -\alpha,\ \nu = -\lambda$; ν Schaden-Supremum
- $\xi = 0$: $\sigma = \sigma^*$; Exponential

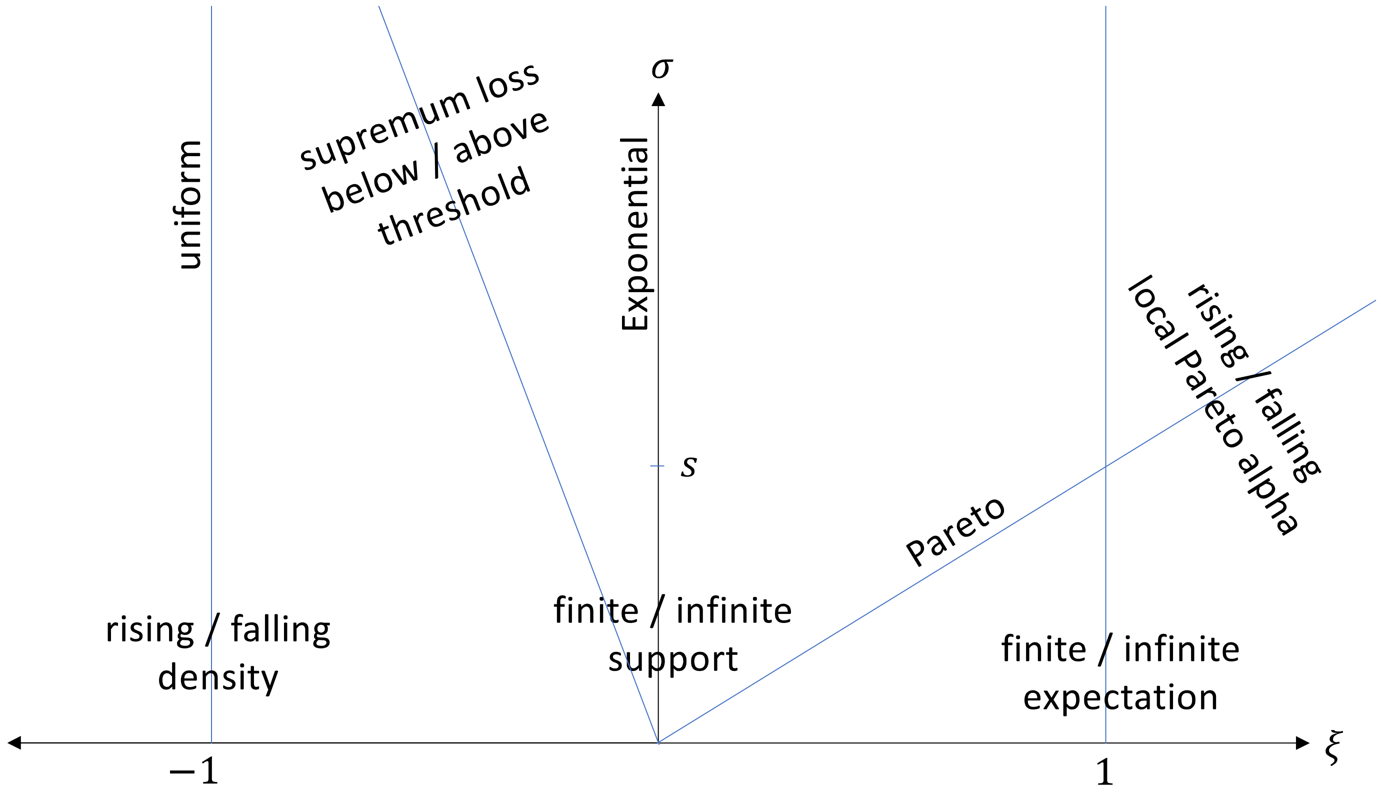
$$P(Z > x | Z > s) = \left(\frac{\xi s + \sigma^*}{(\xi x + \sigma^*)^+}\right)^{1/\xi} =$$

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

Lokales Pareto-Alpha

- Approximiert Vtl.-Fkt $F(x)$ lokal mit Paretokurve
- Allg. Definition: $\alpha_d := d\,\frac{f(d)}{1-F(d)}$
- für die GPD $\alpha_d = \frac{d}{\sigma+\xi(d-s)}$
- für $\xi > 0$ $\alpha_d = \alpha\,\frac{d}{d+\lambda}$

Karte der GPD-Parameter



Drei-Layer-Probleme

Gegebener Input:

- Risikoprämien für 3 Referenz-Layer
- Frequenzen für 3 Thresholds
- Mischung aus beiden

Thresholds sind Grenzwerte von Layern:
Frequenz bei Threshold = *risk rate on line* eines infinitesimal dünnen Layers

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

Beispiel KH

Formuliert als (gemischtes) Drei-Layer-Problem:

- Layer **2 xs 1**: *RRoL* = **52%**
- Layer **5 xs 5**: *RRoL* = **4,8%**
- Threshold **20**: *Freq.* = **0,5%**

Theorem

Für 3 nicht überlappende Layer mit gegebenen RRoL's

$$r_1 > r_2 > r_3 > 0$$

ist das 3-Layer-Problem (meistens) lösbar:

mit einer **eindeutigen** GPD-Tailverteilung

und einer (eindeutigen) Frequenz f bei der Priorität s des niedrigsten Layers.

Theorem

- Gilt auch für Thresholds und für gemischten Input
- $s = 0$ ist möglich (Ground-up Modell)
- Der TopLayer darf unlimitiert sein ($r_3 = 0$)
- Die Layer 1 and 2 dürfen sich (ein wenig) überschneiden
- Eindeutigkeit gilt für weitere Situationen

Einziger Fall ohne Lösung:

- Layer 1 ist kein Threshold, Layer 3 is limitiert
- $r_1 \gg r_2 \approx r_3$

Beispiel KH

$s = 1$ (Mio. USD)

- $f = 1{,}09$
- $\xi = 0{,}41$ $\qquad\qquad (\alpha = 2{,}44)$
- $\sigma = 0{,}96$

Anmerkungen

- Lösung numerisch leicht zu finden:
  Suche 3 Parameter für 3 RRoL-Gleichungen
- Spezialfall *layer-endpoints problem*:
  Ein Layer with 3 Inputs: Risikoprämie, Frequenzen bei Priorität und Plafond
- Pareto löst analoge **2-Layer-Probleme**
- GPD löst viele **4-Layer-Probleme** mit realistischen Inputs approximativ

Anmerkungen

Das Paper gibt **Modellbildungs-Rezepte** für diverse Situationen mit wenig Dateninput

- Stückweise GPD löst ***n*-Layer-Probleme** exakt

Vorgehensweise:

- GPD für die höchsten 3 (notfalls nur 2) Layer
- Füge untere Layer sukzessive an, jeweils mittels Lösung eines *layer-endpoints problem*

Beispiel KH, Variante

- 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%**

Approximativer GPD-Fit:

- 3 Parameter, 4 Gleichungen und eine Abstands-Metrik, z.B. Summe der quadrierten RRoL-Differenzen (absolut, relativ, oder dazwischen)

Beispiel KH, Variante

Exakter Stückweise-GPD-Fit:

- GPD für die Layer 2 – 4
- Resultierende Schadenfrequenz bei 3 is 19%; das ist kompatibel mit dem RRoL 52% von Layer 1
- Wähle kompatible Frequenz bei 1 (>52%) und löse das *layer-endpoints problem* für Layer 1

Modellrisiko

… ist hoch bei wenig Daten, aber:

- **Hauptunsicherheit** ist der Erwartungswert des Layer-Gesamtschadens – sowie manchmal das Schadenzahlmodell
- Höhere Momente des Layer-Gesamtschadens sind wenig unsicher, v.a. für mittlere Layer
- GPD ist eine *Wahl*, aber eine gute, in **praktischer** und **statistischer** Hinsicht: andere Verteilungen würden ähnlichen Output liefern, sind aber weniger praktisch

Parameterfreie Ungleichung

Limitierter Layer: Haftung c, Layerschaden X,

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

mit Schadenfreq. f, Totalschadenfreq. g, RRoL r:

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

- Intervall ist klein bei heavy-tailed SHV
- engeres Intervall for konkave Vtl.-Fkt.
- analoge Intervalle für höhere Momente

Parameterfreie Ungleichung

Entsprechendes Intervall für den Layer-Gesamtschaden S folgt aus

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

Mit der *Contagion* der Schadenzahl

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

einer sehr nützlichen Größe

Zusammenfassung

Bei Extrapolation

- unter den niedrigsten oder
- über den höchsten

Referenzlayer hängen die Ergebnisse stark von der Wahl der parametrischen Verteilung ab.

Für alle Layer innerhalb der Zone der Referenzlayer ist das Modellrisiko sehr niedrig.

Fazit

Die Modellbildung durch Lösung des Drei-Layer-Problems ist ein **mächtiges** Werkzeug und bei knapper Datenlage ein sehr guter **Trade-off** zwischen dem *statistischen Anspruch* und der *Notwendigkeit, Ergebnisse zu liefern.*

Danke für‘s Zuhören.

Details stehen im Paper (ASTIN Colloquium 2021 Website, Updates auf SSRN).

michael_fackler@web.de