2022 ASTIN Colloquium

Tuesday June 21, Parallel Session 3

Thinning of Loss Counts and the Mixed Contagion Model

Dr Michael Fackler

Independent Consulting Actuary,

Munich, Germany

About the speaker

- Dr. (rer. nat.) Michael Fackler, Munich, Germany
- 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 consulting actuary
- Specialized in: non-life reinsurance pricing, dealing with scarce data

Abstract

We investigate representations of loss count distributions that are largely invariant to **thinning**: where most parameters do not change when we, e.g., go from ground-up losses to the subset of losses exceeding a certain threshold. An important thinning-invariant parameter is the **contagion**, which was introduced for the Panjer (a,b,0) class of distributions, but can be defined in general.

We show that one can find loss count representations where all **parameters** are **thinning-invariant** but one. Such representations help compare heterogeneous data sets, like loss records having different reporting thresholds.

As an example, we study the **Mixed Contagion model**, which strongly generalizes the common Poisson-Gamma loss count model, by allowing both for global and local fluctuations. We estimate its parameters by **combining** two data sources.

Outline

- Thinning
- Contagion
- Mixed Contagion model

- Appendix: formulary

Motivation

Can we infer parameters of a collective of risks from a subset, or vice versa? E.g.

- pool vs portfolio
- all losses vs subclass

Can we combine data sets of different kind for inference? E.g.

- all losses of a small portfolio
- the million dollar losses of a large portfolio

Thinning

Relates losses of a portfolio to a subset, according to some **selection criterion**, e.g.

- loss size
- loss cause
- MTPL: containing personal injury component (not only material damage)

Formal setting

Let B_i be the Bernoulli variable describing whether the i-th loss belongs to the subset of interest

- Thinned loss count

$$\bar{N} = \sum_{i=1}^{N} B_i$$

Generalized collective model:

- The B_i constitute a collective model (but the loss sizes X_i maybe not)

Contagion

Suppose N has finite 1st and 2nd moment.

With $\lambda = \mathrm{E}(N)$ we define the *contagion* of N

$$c = \mathrm{Ct}(N) := \left(\frac{\mathrm{Var}(N)}{\mathrm{E}(N)} - 1\right)\frac{1}{\mathrm{E}(N)}$$

We have $$\mathrm{Var}(N) = \lambda + c\lambda^2$$

In the generalized collective model the contagion is **thinning-invariant**; with $\bar{\lambda} = E(\bar{N})$ we have the **same polynomial**

$$\mathrm{Var}(\bar{N}) = \bar{\lambda} + c\bar{\lambda}^2$$

Looking deeper

- Analogous polynomial representations exist for higher central and raw moments (if finite).
- If a loss count distribution can be parameterized by its first n moments, this can be converted to a representation where all parameters but λ are invariant to thinning.
- Key idea: The normalized factorial moments $\varphi_k := \mathrm{E}\left(N_{(k)}\right)/\lambda^k$ are thinning-invariant.
- In particular, $\mathrm{Ct}(N) = \varphi_2 - 1$

Pooling NegBin risks

Consider v Poisson-Gamma loss count variables U_i, having each expectation θ, constituting a portfolio

- $N = \sum U_i$ has expectation $\lambda = v\theta$

Case 1) $U_i \sim Poi(\theta Q)$ with **common** mixing Gamma (global fluctuation), $\text{Var}(Q) = \beta$

- $\text{Var}(U_i) = \theta + \beta\theta^2$, $\text{Var}(N) = \lambda + \beta\lambda^2$, $\text{Ct}(N) = \beta$

Case 2) $U_i \sim Poi(\theta R_i)$ with **iid** mixing Gammas (local fluctuation), $\text{Var}(R_i) = \kappa$

- $\text{Var}(U_i) = \theta + \kappa\theta^2$, $\text{Var}(N) = \lambda + \frac{\kappa}{v}\lambda^2$, $\text{Ct}(N) = \frac{\kappa}{v}$

Mixed Contagion model

- Loss count N of portfolio having volume v has expectation $v\theta$ and contagion $\beta + \frac{\kappa}{v}$
- $\mathrm{Var}(N) = v(\theta + \theta^2\kappa) + v^2\theta^2\beta$
- Class of models, specifying 2 moments via 3 parameters
- $\kappa < 0$ possible (to some extent)
- Embraces and generalizes the two above NegBin cases, intermediate situations possible
- E.g. $U_i \sim Poi(\theta R_i Q)$: global + local fluctuation

Parametric examples

- **Poisson-2xGamma**: If Q and the R_i are Gamma distributed, we get a *mixed NegBin* model, which is similar to NegBin, but not as handy
- Analogous **Poisson-2xLognormal** yields *Poisson-Lognormal*, which is (numerically) inconvenient, but used in biostatistics

Parameter-free alternative

- Idea: Don't specify the MC distribution further, work with empirical first and second moment
- Approach analogous to the moment estimators for NegBin (see Th. Mack's book)
- For estimating three parameters, one needs to combine two data sets, which must be (possibly thinned) typical parts of the same pool of risks
- Long data history needed for empirical variance

Mathematical core

- In the generalized collective model with MC loss count, for a typical sub-portfolio having volume v,

$$\mathrm{E}(N) = v\theta, \quad \mathrm{Ct}(N) = \beta + \frac{\kappa}{v}$$

- and after thinning

$$\mathrm{E}(\bar{N}) = v\bar{\theta}, \quad \mathrm{Ct}(\bar{N}) = \beta + \frac{\kappa}{v}$$

Procedure

1st data: volumes v_i, (possibly thinned) loss counts $\bar{n}_i$

- The empir. frequencies $t_i = \frac{\bar{n}_i}{v_i}$ have expectation $\bar{\theta}$
- Choose weights w_i, calculate weighted sample mean m (**estimates** $\bar{\theta}$) and respective variance s.
- Then $\frac{s}{m^2} - \frac{b}{m}$ **estimates** $a\beta + b\kappa$

where $a = \sum\left(w_i - w_i^2\right)$, $b = \sum \frac{w_i - w_i^2}{v_i}$

Procedure

2nd data set: possibly different observation period, (possibly otherwise thinned) loss counts $\dot{n}_j'$, quite different volumes v_j',

- As before: calculate the t_j'; choose weights w_j'; calculate m' (**estimates** $\dot{\theta}$) and s', a' and b'
- Then $\frac{s'}{m'^2} - \frac{b'}{m'}$ **estimates** $a'\beta + b'\kappa$

Now solve **system of equations** for β and κ

- Calculate (estimated) variances, **reiterate** with weights inversely proportional to variances

Thanks

For details see paper, which will be updated on SSRN.com

For more on contagion and Mixed Contagion model see Chapter 5 of

Fackler M (2017) *Experience rating of (re)insurance premiums under uncertainty about past inflation.* PhD thesis, Universität Oldenburg, urn:nbn:de:gbv:715-oops-33087

WANTED: loss count data sets for application

michael_fackler@web.de

Appendix: Moments

- raw $m_k = \mathrm{E}\left(N^k\right), \;\; \lambda = m_1 = \mathrm{E}(N)$
- central $\mu_k = \mathrm{E}\left((N-\lambda)^k\right)$
- factorial $\mu_{(k)} = \mathrm{E}\left(N_{(k)}\right) = P_N^{(k)}(1_-)$ (pgf P_N)

and analogously for $\bar{N}$

In the generalized collective model we have with $p = \mathrm{E}(B_i)$: $\bar{\lambda} = p\lambda$,

$$\bar{\mu}_{(k)} = p^k \mu_{(k)}, \qquad \bar{\mu}_{(k)}/\bar{\lambda}^k = \mu_{(k)}/\lambda^k =: \varphi_k$$

Polynomials

$$\mu_{(k)} = \varphi_k \lambda^k$$

$$m_k = \sum_{i=1}^{k} \left\{ {k \atop i} \right\} \mu_{(i)} = \lambda + \sum_{i=2}^{k} \gamma_{ki} \lambda^i$$

where $\gamma_{ki} = \left\{ {k \atop i} \right\} \varphi_i$, $\left\{ {k \atop i} \right\}$ Stirling no. of 2nd kind

$$\mu_k = \sum_{i=0}^{k} \binom{k}{i} m_{k-i} (-\lambda)^i = \ldots = \lambda + \sum_{i=2}^{k} \varepsilon_{ki} \lambda^i$$

with linear combinations ε_{ki} of the φ_j

- For $\bar{N}$ same polynomials with $\bar{\lambda}$ instead of λ

Parameter estimate for MC model

$$t_i = \frac{\bar{n}_i}{v_i},\quad m = \sum w_i t_i,\quad s = \sum w_i (t_i - m)^2$$

$$\mathrm{E}(\bar{N}_i) = v_i \bar{\theta},\quad \mathrm{Var}(\bar{N}_i) = v_i(\bar{\theta} + \bar{\theta}^2 \kappa) + v_i^2 \bar{\theta}^2 \beta$$

Set $a = \sum \left(w_i - w_i^2\right),\quad b = \sum \frac{w_i - w_i^2}{v_i}$

- m is an unbiased estimator for $\bar{\theta}$,
  s for $b\bar{\theta} + b\bar{\theta}^2\kappa + a\bar{\theta}^2\beta$
- $\frac{s}{m^2} - \frac{b}{m}$ is a natural estimator for $a\beta + b\kappa$