ETH zürich

Insurance Data Science Conference

Selim Gatti, Risklab, ETH Zürich

18 June 2024, Stockholm

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Project**

- My presentation will be about a joint project with Prof. Dr. **Mario Wüthrich**.
- This project led to a preprint available on Arxiv under the name *Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class*.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Problem statement**

- Denote by X the total financial loss faced by an insured. After subtracting a deductible $d > 0$ and applying a maximal cover of size $M > 0$, the insurance claim Y paid by the insurer is given by the random variable

$$Y = \min\{(X-d)_+,\, M\} \mid X > d.$$

  We say, the total financial loss is *lower-truncated* at $d > 0$ and *right-censored* at $M > 0$.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Problem statement**

- When modeling insurance claims, the total financial loss X is typically assumed to be an absolutely continuous random variable on $(0, \infty)$. In this case, the insurance claim Y has as density

$$f_Y(z) = \frac{f_X(z+d)}{1 - F_X(d)},$$

for $z \in [0, M)$ and a point mass p in M being equal to

$$p = \frac{1 - F_X(d+M)}{1 - F_X(d)}.$$

- The log-likelihood function of the random variable Y for an unknown parameter θ is then given by

$$\theta \mapsto \ell_Y(\theta) = \log\left(f_Y(\theta)\right) \mathbb{1}_{\{Y<M\}} + \log\left(p(\theta)\right) \mathbb{1}_{\{Y=M\}}.$$

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Problem statement**

- Fitting such a model with Maximum Likelihood Estimation (MLE) can be difficult because we need an analytically tractable form for both the density f_X and its distribution function F_X.
- This is not the case for most statistical models that can fit general non-life insurance claims, e.g. the gamma or the log-normal model.
- In such cases, one can either rely on numerical integration of the density (which can be computationally demanding) or use a version of the Expected Maximization algorithm (which has its drawbacks).

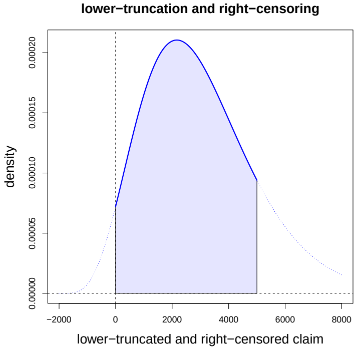

Figure: Lower-truncated and right-censored claim with $d = 2000$ and $M = 5000$.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **MBBEFD class**

- We take another approach in this project by introducing a class of models that can fit lower-truncated and right-censored random variables with fully tractable distribution functions, absolutely continuous densities, point masses and means, i.e. they are all of closed form.
- For this, we start from the MBBEFD class of distributions introduced by Bernegger in 1997 [1] in the reinsurance literature.
- We show that this class contains unimodal densities which are symmetric around their mode.
- However, to model general non-life insurance claims, we rather need unimodal and skewed densities.
- Therefore, we extend the MBBEFD class to a larger class of distributions including monotonically decreasing, unimodal and monotonically increasing densities. Our extension provides new families of lower-truncated and right-censored random variables allowing for skewness in the absolutely continuous part of the density. We call this new class of distributions *the Bernegger class*.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Bernegger class**

- We show that without loss of generality, we can work with normalized losses lying on the unit interval $[0, 1]$.
- To introduce this new family, we start from the *exposure curve* of a normalized loss $Z \sim F_Z$, which is defined by

$$z \;\mapsto\; G(z) = \frac{\int_0^z 1 - F_Z(s)\, ds}{\int_0^1 1 - F_Z(s)\, ds} = \frac{\int_0^z 1 - F_Z(s)\, ds}{\mathbb{E}[Z]},$$

  for $z \in [0, 1]$.
- This exposure curve $G : [0, 1] \to \mathbb{R}$ is a non-decreasing, concave, and twice continuously differentiable function with $G(0) = 0$, $G(1) = 1$, and $G'(0) > 0$.
- As formalized in the next theorem, it turns out that any function satisfying these properties allows to derive the distribution of a random variable taking values in $[0, 1]$, with an absolutely continuous density on $[0, 1)$ and a point mass in 1.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Bernegger class**

Theorem

Let $G : [0, 1] \to \mathbb{R}$ be a non-decreasing, concave, and twice continuously differentiable function with $G(0) = 0$, $G(1) = 1$, and $G'(0) > 0$. The function $F_Z : [0, 1] \to \mathbb{R}$ defined by

$$F_Z(z) = \left(1 - \frac{G'(z)}{G'(0)}\right) \mathbb{1}_{\{z<1\}} + \mathbb{1}_{\{z=1\}}$$

is a distribution function on $[0, 1]$. Furthermore, this distribution has as density

$$f_Z(z) = -\frac{G''(z)}{G'(0)},$$

for $z \in [0, 1)$, and a point mass in 1 given by

$$p = \frac{G'(1)}{G'(0)}.$$

Finally, the mean of $Z \sim F_Z$ is equal to $\mathbb{E}[Z] = 1/G'(0)$.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Bernegger class**

- In order to define members of the Bernegger class, we start from a general class of exposure curves
$$z \mapsto G(z) = \frac{h(b(z)) - h(b(0))}{h(b(1)) - h(b(0))},$$
where the function h is called the *link function*, whereas the function b is called the *inner function*. These functions have to be carefully chosen in order to obtain a well-defined exposure curve.
- Bernegger's original choice was $b(z) = \alpha + \beta^z$ and $h(x) = \log(x)$ for carefully chosen parameters α and β.
- We state in our paper the conditions on the inner function in order to derive well-defined exposures curves for a logarithmic or exponential link function.
- Interestingly, lower-truncated and right-censored **exponential, logistic and Pareto distributions** are members of the Bernegger class.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Flexibility of the Bernegger class**

- The Bernegger class offers a high flexibility. By using so-called one-inflated distributions, one can add more flexibility to the point mass at the right end.
- Indeed, suppose that Z belongs to the Bernegger class. If we first consider the conditional density of the random variable $Z_0 \stackrel{(\mathrm{d})}{=} Z|_{\{Z<1\}}$ and then add a flexible point mass q in 1, the associated distribution of the new random variable again belongs to the Bernegger class.
- Similarly, one could in principle add a point mass at the left end allowing to model left- and right-censored random variables.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Real dataset**

- Using a real dataset consisting of $126'026$ private insurance claims, we perform Maximum Likelihood Estimation and the best fit is obtained by a member of the Bernegger class that uses an exponential link function. The fit on the absolutely continuous part seems adequate and this model manages to match the empirical point mass in 1 being equal to 0.034.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Real dataset**

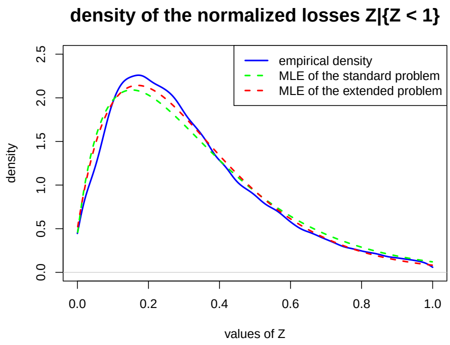

Figure: Power exponentially linked exposure example: densities of the random variable $Z|_{\{Z<1\}}$.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Real dataset**

	Point mass	Mean	$\ell_{Z\|_{\{Z<1\}}}$	ℓ_Z	AIC
Empirical density (Blue)	0.034	0.339	-	-	-
Standard MLE density (Green)	0.025	0.341	34 068	15 199	-30 390
Flexible MLE density (Red)	0.034	0.339	34 402	15 731	-31 453

Table: Power exponentially linked exposure example: results.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Changes in the contract**

- The insurer might be interested in understanding how a change in the deductible and/or maximal cover affects the expected claim size.
- We discuss the cases where this is possible and we point out that in general, extrapolating the observed claims density leads to several possible candidates for the original distribution of the total financial loss.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Extrapolation**

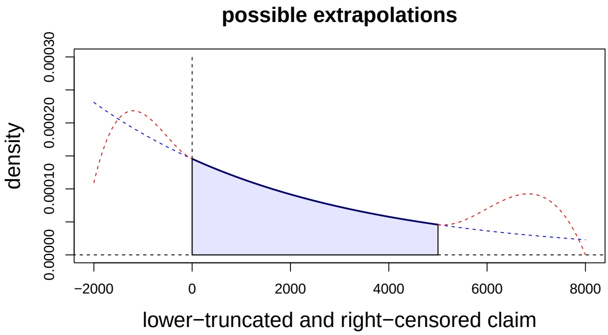

Figure: Two possible extrapolations of the density of a lower-truncated and right-censored exponential random variable with $d = 2000$ and $M = 5000$.

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Closedness of the Bernegger class**

- Finally, by looking at this problem, we noticed that the Bernegger class is closed under lower-truncation and right-censoring as formalized in the next proposition.

Proposition

Let Z be a member of the Bernegger class with exposure curve G and let $0 \leq \bar{d}, \bar{M} \leq 1$ such that $\bar{d} + \bar{M} \leq 1$. Moreover, define the scaled lower-truncated and right-censored random variable

$$\tilde{Z} = \frac{1}{\bar{M}} \min\{(Z - \bar{d})_+, \bar{M}\} \,|\, Z > \bar{d}.$$

This random variable belongs again to the Bernegger class and its exposure curve is given by

$$\tilde{G} : [0,1] \to [0,1], \; z \mapsto \frac{G(\bar{d} + z\bar{M})}{G(\bar{d} + \bar{M})}.$$

Modeling lower-truncated and right-censored insurance claims with an extension of the MBBEFD class - **Next steps**

- A potential next step is to try to further classify the members of the Bernegger class and to understand more precisely the role of the link function on the resulting distributions.
- Another next step is to lift the Bernegger class to regression models such as generalized linear models, study model fitting and analyze the properties of the parameter estimates such as consistency and asymptotic normality of the fitted parameters.

ETH zürich

Do you have any questions?

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Risklab
Selim Gatti

References

S. Bernegger, "The Swiss Re exposure curves and the MBBEFD distribution class," *ASTIN Bulletin*, vol. 27, no. 1, pp. 99–111, 1997.