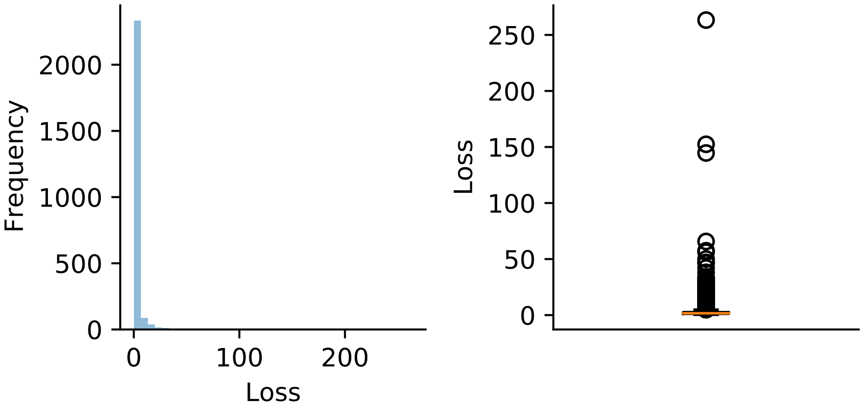
Sequential Monte Carlo samplers to fit and compare insurance loss models

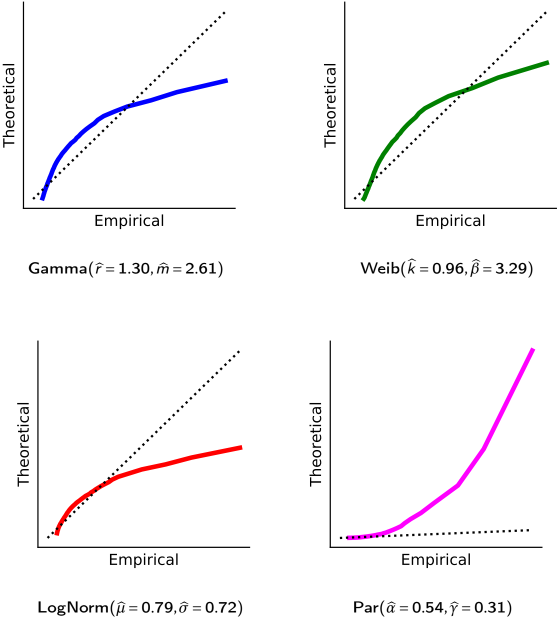
Pierre-O. Goffard

Institut de Science Financières et d'Assurances
pierre-olivier.goffard@univ-lyon1.fr

13 juin 2021



Empirical distribution of the danish fire insurance losses.


Theoretical
Empirical
Gamma($\hat{r} = 1.30, \hat{m} = 2.61$)
Theoretical
Empirical
Weib($\hat{k} = 0.96, \hat{\beta} = 3.29$)
Theoretical
Empirical
LogNorm($\hat{\mu} = 0.79, \hat{\sigma} = 0.72$)
Theoretical
Empirical
Par($\hat{\alpha} = 0.54, \hat{\gamma} = 0.31$)

Composite models

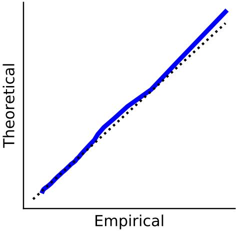
The **pdf** of a composite model is defined as

$$f(x) = \begin{cases} p\frac{f_1(x)}{F_1(\gamma)}, \text{ si } x \leq \gamma, \\ (1-p)\frac{f_2(x)}{1-F_2(\gamma)}, \text{ si } x > \gamma, \end{cases}$$

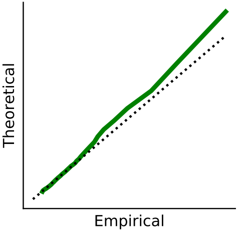
where f_1, F_1, f_2, and F_2 are the **pdf** and (**cdf**) of the models for the belly and the tail of the loss distribution respectively.

- $f_1 =$ gamma, Weibull or lognormal
- $f_2 =$ Pareto

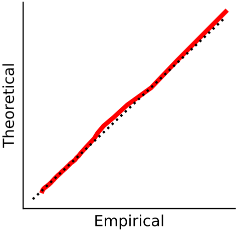
Y. Wang, I. H. Haff, and A. Huseby, "Modelling extreme claims via composite models and threshold selection methods," *Insurance : Mathematics and Economics*, vol. 91, pp. 257–268, mar 2020.



Gamma($\widehat{r} = 35.68$) – Par($\widehat{\alpha} = 1.31, \widehat{\gamma} = 1.16$)



Weib($\widehat{k} = 14.03$) – Par($\widehat{\alpha} = 1.26, \widehat{\gamma} = 1.00$)



LogNorm($\widehat{\sigma} = 0.19$) – Par($\widehat{\alpha} = 1.32, \widehat{\gamma} = 1.21$)

Bayesian statistics

Let $\mathscr{M}$ be a model with parameter θ, and $\mathbf{x}$ some observed data.

- Bayesian statistics targets the posterior distribution of the parameter

$$\pi(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta)\pi(\theta)}{\int_\Theta p(\mathbf{x}|\theta)\pi(\theta)\mathrm{d}\theta} = \frac{p(\mathbf{x}|\theta)\pi(\theta)}{Z(\mathbf{x})},$$

by updating the prior $\pi(\theta)$ via the likelihood $p(\mathbf{x}|\theta)$.

↪ Model calibration

If many models $\mathscr{M}_1,\ldots,\mathscr{M}_K$ are competing

- The posterior model evidence of each model follow from

$$\pi(M_i|\mathbf{x}) = \frac{p(\mathbf{x}|M_i)\pi(\mathscr{M}_i)}{\sum_{j=1}^K p(\mathbf{x}|M_j)\pi(\mathscr{M}_j)},\ i = 1,\ldots,K.$$

↪ Select or combine models

Sequential Monte Carlo Sampler

Let $\pi_s,\ s=0,\ldots,t$ be a sequence of intermediary distributions such that $\pi=\pi_0$ and $\pi_t=\pi(\cdot|\mathbf{x})$ represented by particle clouds $(W_i^s,\theta_i^s),\ i=1,\ldots,N$.

1. **Reweight** :

$$W_i^{s+1}\propto w_i^{s+1}=\frac{\pi_{s+1}(\theta_i^s)}{\pi_s(\theta_i^s)} \text{ such that } ESS>\rho\cdot N$$

2. **Select** :

$$(\widetilde{\theta}_1^{s+1}\ldots,\widetilde{\theta}_N^{s+1})\sim\left\{(W_1^{s+1},\theta_1^s)\ldots,(W_N^{s+1},\theta_1^s)\right\}$$

3. **Move** :

$$\theta_i^{s+1}\sim K_H(\widetilde{\theta}_i^{s+1},\cdot) \text{ and } W_i^{s+1}\leftarrow 1/N \text{ for } i=1,\ldots,N$$

We have

$$\theta_1^t,\ldots,\theta_N^t\sim\pi(\cdot|\mathbf{x}), \text{ and } Z(\mathbf{x})\approx\prod_{s=1}^{t}\left(\frac{1}{N}\sum_{i=1}^{N}{w_i}^s\right)$$

. P. D. Moral, A. Doucet, and A. Jasra, "Sequential monte carlo samplers," *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, vol. 68, pp. 411–436, jun 2006.

How to define these intermediary distributions?

- Simulated annealing

$$\pi_s(\theta) \propto \pi(\theta|\mathbf{x})^{\tau_s}\pi(\theta),$$

  where $0 = \tau_0 < \ldots < \tau_t = 1$.

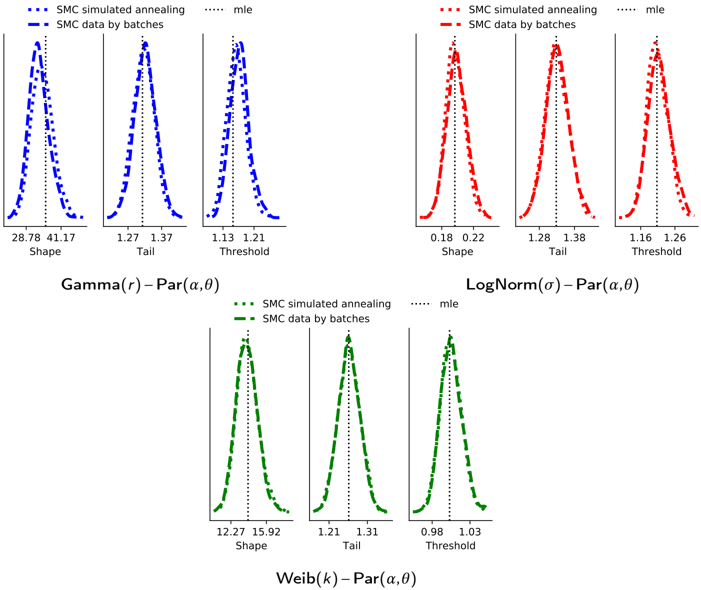
- Data by batches

$$\pi_s(\theta) \propto \pi(\theta|x_1, \ldots, x_{n_s})\pi(\theta)$$

  where $0 = n_0 < \ldots < n_t = n$.

. R. M. Neal, "Annealed importance sampling," *Statistics and Computing*, vol. 11, no. 2, pp. 125–139, 2001.

. N. Chopin, "A sequential particle filter method for static models," *Biometrika*, vol. 89, pp. 539–552, aug 2002.



SMC simulated annealing
SMC data by batches
mle
28.78
41.17
Shape
1.27
1.37
Tail
1.13
1.21
Threshold
Gamma(r) − Par(α,θ)
SMC simulated annealing
SMC data by batches
mle
0.18
0.22
Shape
1.28
1.38
Tail
1.16
1.26
Threshold
LogNorm(σ) − Par(α,θ)
SMC simulated annealing
SMC data by batches
mle
12.27
15.92
Shape
1.21
1.31
Tail
0.98
1.03
Threshold
Weib(k) − Par(α,θ)

Inference summary

Methods	Models	$Z(\mathbf{x})$	$\pi(m\|\mathbf{x})$	Time [1]
Simulated	lnorm-par	-3884.39	0.00	$423s$
Annealing	wei-par	-3855.92	1.00	
	gam-par	-3877.01	0.00	
Data	lnorm-par	-3882.52	0.00	$2065s$
by	wei-par	-3858.24	1.00	
batch	gam-par	-3878.15	0.00	

1. $N = 2000$ paralell on a 40 cores server

Conclusions and perspectives

Implementation of the SMC sampler to fit and compare composite models

⚠ Lack of fit of the Pareto tail ⇒ alternative models for the tail

B. Grün and T. Miljkovic, "Extending composite loss models using a general framework of advanced computational tools," *Scandinavian Actuarial Journal*, vol. 2019, pp. 642–660, apr 2019.

⚠ Better estimate of the higher order quantiles ⇒ minimum distance estimator

E. Bernton, P. E. Jacob, M. Gerber, and C. P. Robert, "On parameter estimation with the Wasserstein distance," *Information and Inference : A Journal of the IMA*, vol. 8, pp. 657–676, oct 2019.