Estimation of extreme quantiles from heavy-tailed distributions in a semi-parametric location-dispersion regression model

Communication at the 4th "Rencontres des jeunes chercheurs africains en France"

by

Aboubacrène Ag AHMAD[1,4]

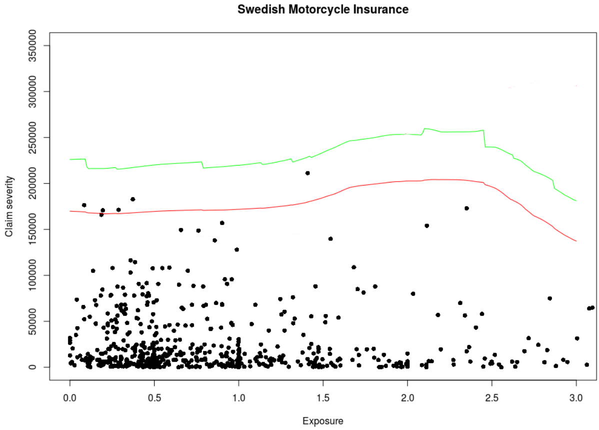
in collaboration with

Stéphane GIRARD[2], Antoine USSEGLIO-CARLEVE[3], Aliou DIOP[4] *and* El Hadji DEME[4]

[1] Université des Sciences, des Techniques et des Technologies de Bamako, FST.
[2] Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK.
[3] Univ. Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes.
[4] LERSTAD, Université Gaston Berger de Saint-Louis.

Motivation

☞ Estimation of extreme conditional quantiles: a recurrent statistical problem in the modeling of extreme events in finance, insurance, epidemiology, environmental science, ...



Semi-parametric location-dispersion regression model

Consider the following regression model between a random response variable $Y \in \mathbb{R}$ and a deterministic covariate vector $x \in \Pi \subset \mathbb{R}^d$, $d \geq 1$:

$$Y = a(x) + b(x)\,Z,$$

where

- $a : \Pi \to \mathbb{R}$: regression function (unknown);
- $b : \Pi \to \mathbb{R}^+ \setminus \{0\}$: dispersion function (unknown);
- $Z \in \mathbb{R}$: is a heavy-tailed random variable with tail-index $\gamma > 0$.

Inference

Step 1: Estimation of the regression and dispersion functions

☞ Kernel estimators of $a(\cdot)$ and $b(\cdot)$: $\hat{a}_n$ and $\hat{b}_n$.

Step 2: Estimation of the tail-index

☞ The tail-index γ is estimated using the standard extreme procedures from the residuals given by:

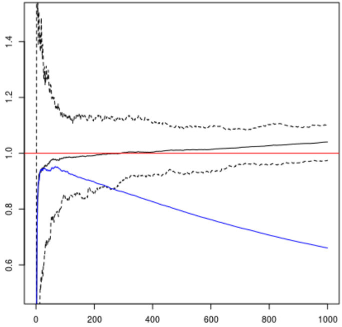
$$\hat{Z}_i = (Y_i - \hat{a}_n(x_i))/\hat{b}_n(x_i),\ i = 1, \ldots, n.$$

Step 3: Estimation of extreme conditional quantiles

☞ Plug-in estimator built from the residuals.

Advantages of the semi-parametric model

- Reduction of bias compared to the classical estimation approaches by using the residuals Z_i instead of the response variables Y_i.



- No curse of dimensionality compared to the purely nonparametric approaches.