Extremile regression

Gilles Stupfler (ENSAI & CREST)

Joint work with Abdelaati Daouia (Toulouse School of Economics)
and Irène Gijbels (KU Leuven)

Insurance Data Science Conference, 18th June 2021

ENSAI École nationale de la statistique et de l'analyse de l'information

CREST

Any quantile is a median, any extremile is a mean

Focus of this presentation

Study a class of L^2-based risk measures which are comonotonically additive and expectations over the whole of the distribution.

Assume that F is continuous and strictly increasing. Then the quantile q_τ of F satisfies

$$[F(q_\tau)]^{\log(1/2)/\log(\tau)} = \tau^{\log(1/2)/\log(\tau)} = \exp(\log(1/2)) = 1/2.$$

So q_τ is the median of Z_τ, where Z_τ has c.d.f. $z \mapsto [F(z)]^{\log(1/2)/\log(\tau)}$.

Let $r(\tau) = \log(1/2)/\log(\tau)$ and $K_\tau(t) = t^{r(\tau)}$.

Definition (Extremile)

The extremile of order τ of Y is the expectation of $Z_\tau = Z_\tau(Y)$:

$$\xi_\tau = \mathbb{E}(Z_\tau), \text{ with } Z_\tau \text{ having c.d.f. } K_\tau \circ F.$$

Alternative formulations of extremiles

When $r(\tau)$ is a positive integer, we then have

$$\xi_\tau = \mathbb{E}(\max(Y_1, \ldots, Y_{r(\tau)}))$$

where the Y_i are independent copies of Y.

Extremiles are coherent and comonotonically additive. They are not elicitable, but their formulation as a minimizer suggests they can be backtested with a natural methodology (work in progress).

For regression: If $J_\tau(t) = K'_\tau(t)$, it holds that

$$\xi_\tau = \arg\min_{\theta \in \mathbb{R}} \mathbb{E}(J_\tau(F(Y))[(Y-\theta)^2 - Y^2]).$$

So just like expectiles, extremiles are defined through an asymmetric least squares criterion, but with a different weighting scheme.

Extremile regression

Let $(Y, X) \in \mathbb{R} \times \mathbb{R}^d$ be a response-covariate pair. Assume that the conditional c.d.f. $F(\cdot|x)$ is continuous. The τth extremile of this c.d.f. defines the τth regression extremile of Y given $X = x$.

Definition (Regression extremile)

The τth regression extremile of Y given $X = x$ is

$$\xi_\tau(x) = \arg\min_{\theta \in \mathbb{R}} \mathbb{E}(J_\tau(F(Y|x))[(Y - \theta)^2 - Y^2] \,|\, X = x).$$

This definition requires $\mathbb{E}(|Y| \,|\, X = x) < \infty$ to make sense.

Regression extremiles keep the interpretation of extremiles (expectation of maxima...) but applied to the conditional distribution instead.

Assume that a random sample (Y_i, X_i), $1 \le i \le n$ is available.

Estimation - extreme case

Focus now on the case $\tau = \tau_n \uparrow 1$ as $n \to \infty$. An extrapolation method is needed! Assume that

$$\forall y > 0, \lim_{t\to\infty} \frac{q_{1-(ty)^{-1}}(x)}{q_{1-t^{-1}}(x)} = y^{\gamma(x)}.$$

The conditional distribution has a Pareto-type tail with index $\gamma(x)$.

In this case, if $\mathbb{E}(\max(-Y,0)\,|\,X=x) < \infty$ and $0 < \gamma(x) < 1$ then

$$\frac{\xi_\tau(x)}{q_\tau(x)} \to \mathcal{G}(\gamma(x)) \text{ as } \tau \uparrow 1,$$

where $\mathcal{G}(s) = \Gamma(1-s)\{\log 2\}^s$ and Γ is Euler's Gamma function.

This means that

$$\widehat{\xi}_{\tau_n}(x) = \mathcal{G}(\widehat{\gamma}(x))\widehat{q}_{\tau_n}(x)$$

is an estimator of the extreme regression extremile $\xi_{\tau_n}(x)$.

Our choices? Focus on $d = 1$ for simplicity. We set first

$$\widehat{F}_{\mathrm{NW}}(y|x) = \sum_{i=1}^{n} \mathbb{1}\{Y_i \le y\} L\left(\frac{x - X_i}{h_n}\right) \Big/ \sum_{i=1}^{n} L\left(\frac{x - X_i}{h_n}\right).$$

Here L is a p.d.f. on $\mathbb{R}$. We then define a kernel estimator of $\widehat{\gamma}(x)$ as

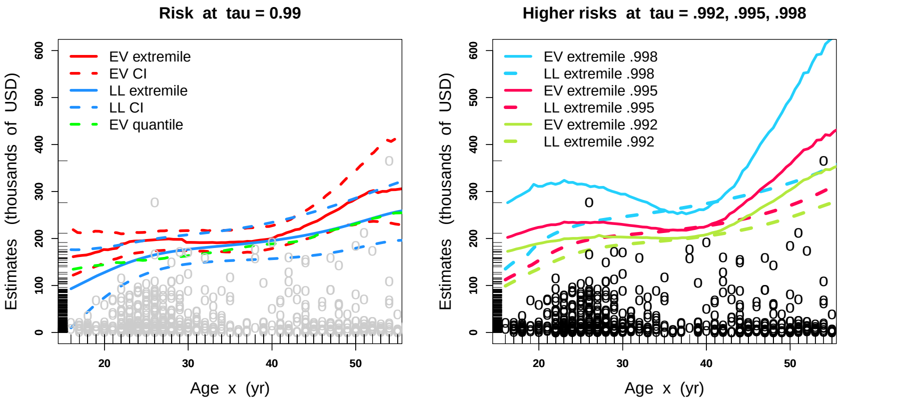
$$\widehat{\gamma}(x) = \frac{\sum_{j=1}^{J}\left[\log \widehat{F}_{\mathrm{NW}}^{-1}(1 - t_j(1 - a_n)|x) - \log \widehat{F}_{\mathrm{NW}}^{-1}(a_n|x)\right]}{\sum_{j=1}^{J} \log(1/t_j)}$$

where $1 = t_1 > t_2 > \cdots > t_J > 0$ are J weights. This is a kernel version of the (generalized) Pickands estimator. The estimator of $q_{\tau_n}(x)$ is then

$$\widehat{q}_{\tau_n}(x) = \left(\frac{1 - \tau_n}{1 - a_n}\right)^{-\widehat{\gamma}(x)} \widehat{F}_{\mathrm{NW}}^{-1}(a_n|x).$$

Here $a_n \uparrow 1$ with $nh_n(1 - a_n) \to \infty$: it is "extreme, but not too much".

For $t_j = 1/j$ the variance V_J is minimal for $J = 9$ with $V_9 \approx 1.25$.



Data set dataOhlsson (R package insuranceData) on $n = 670$ motorcycle-related claims recorded by the Swedish insurer Wasa.

Left: Estimates $\widehat{\xi}_{.99}(x)$ and $\widetilde{\xi}_{\mathrm{LL},.99}(x)$ (local linear estimator), corresponding 95% asymptotic confidence intervals, and $\widehat{q}_{.99}(x)$.

Right: Estimates $\widehat{\xi}_{\tau_n}(x)$ and $\widetilde{\xi}_{\mathrm{LL},\tau_n}(x)$ for $\tau_n = .992, .995, .998$.

Discussion

- An extremile above the mean is the mean of a distribution whose weight has been shifted to the right in a simple way.
- Extremiles are L^2 quantities, have various interpretations and closed forms, and do not rely solely on the tail event.
- They can be estimated at central and extreme levels using local linear estimation and semiparametric extrapolation.

Ongoing work and research perspectives? Forecast evaluation, dependent data (marginal/dynamic estimation)...

For (much!) more, see the following two papers:

Daouia, A., Gijbels, I., Stupfler, G. (2019). Extremiles: A new perspective on asymmetric least squares, *JASA* **114**(527): 1366–1381.

Daouia, A., Gijbels, I., Stupfler, G. (2021). Extremile regression, *JASA*, to appear.