EXTREME LEVEL CURVES OF HEAVY-TAILED DISTRIBUTIONS

Alexandre LEKINA

PhD Student, Team Mistis, INRIA Grenoble-Rhône-Alpes & LJK.
http://mistis.inrialpes.fr/people/lekina

Joint work with A. DAOUIA[1], L. GARDES & S. GIRARD

June 2009

LABORATOIRE JEAN KUNTZMANN
MATHÉMATIQUES APPLIQUÉES • INFORMATIQUE

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
INRIA
centre de recherche GRENOBLE - RHÔNE-ALPES

[1] Toulouse School of Economics (GREMAQ), University of Toulouse.

Goal

- Let (X_i, Y_i), $i = 1, ..., n$ be independent copies of the random pair $(X, Y) \in \mathbb{R}^d \times \mathbb{R}$.
- Y is a variable of interest associated with a covariate information X.
- Estimate for all $x \in \mathbb{R}^d$ fixed and for all $\alpha_n \to 0$, the extreme level curves defined as the graphs of the functions $x \in \mathbb{R}^d \to q(\alpha_n|x) \in \mathbb{R}$ (the conditional extreme quantile of order $(1 - \alpha_n)$) verifying

$$\mathbb{P}(Y > q(\alpha_n|x)|x) = \alpha_n,$$

when the conditional cumulative distribution function of Y given $X = x$ is heavy-tailed with tail index $\gamma(x)$, i.e for all $y > 0$,

$$\bar{F}(y|x) = y^{-1/\gamma(x)}\ell(y|x),$$

Goal

$$\bar{F}(y|x) = y^{-1/\gamma(x)}\ell(y|x),$$

- $\gamma(.)$: "the conditional tail index" is an unknown and positive function of the covariate x.
- $\ell(.|x)$ is a slowly-varying function at infinity, i.e for all $\lambda > 0$,

$$\lim_{y\to\infty} \frac{\ell(\lambda y|x)}{\ell(y|x)} = 1.$$

- $\ell(.|x)$ is normalized and differentiable.

Estimator: the inverse of an estimator of the conditional survival function

$$\hat{q}_n(\alpha_n|x) = \hat{\bar{F}}_n^{\leftarrow}(\alpha_n|x) = \inf\left\{t,\ \hat{\bar{F}}_n(t|x) \le \alpha_n\right\}.$$

Requires to estimate the small tail probability

$$\bar{F}(y_n|x) \text{ when } y_n \to \infty \text{ as } n \to \infty.$$

Kernel estimator of $\bar{F}$, (Collomb (1976))

$$\hat{\bar{F}}_n(y|x) = \frac{\frac{1}{nh_n^d}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)\mathbf{1}\{Y_i \geq y\}}{\frac{1}{nh_n^d}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)} = \frac{\hat{\psi}_n(y,x)}{\hat{g}_n(x)},$$

- The kernel function $K(.)$ is positive, bounded and integrable on a compact support $S \subseteq \mathbb{R}^d$.
- The sequence of window-width $h_n \longrightarrow 0$ as $n \longrightarrow \infty$.
- The function $\hat{g}_n(.)$ is the classical kernel estimator of the point distribution function $g(.)$ of X.
- The function $\hat{\psi}_n(y,x)$ is an estimator of $\psi(y,x) = \bar{F}(y|x)g(x)$.

Kernel estimator, (Collomb (1976))

$$\hat{\bar{F}}_n(y|x) = \frac{\frac{1}{nh_n^d}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)\mathbf{1}\{Y_i \geq y\}}{\frac{1}{nh_n^d}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)} = \frac{\hat{\psi}_n(y,x)}{\hat{g}_n(x)},$$

Remark

$$\frac{\hat{\bar{F}}_n(y_n|x)}{\bar{F}(y_n|x)} - 1 = \left[\frac{\hat{\psi}_n(y_n,x) - \mathbb{E}\left[\hat{\psi}_n(y_n,x)\right]}{\hat{g}_n(x)\psi(y_n,x)}\right] + \left[\frac{\mathbb{E}\left[\hat{\psi}_n(y_n,x)\right] - \psi(y_n,x)}{\hat{g}_n(x)\psi(y_n,x)}\right] - \left[\frac{\hat{g}_n(x) - \mathbb{E}\left[\hat{g}_n(x)\right]}{\hat{g}_n(x)}\right] - \left[\frac{\mathbb{E}\left[\hat{g}_n(x)\right] - g(x)}{\hat{g}_n(x)}\right].$$

- The asymptotic distribution of $\hat{\bar{F}}_n(y_n|x)$ may depend both on the behavior of the random terms $\left[\frac{\hat{\psi}_n(y_n,x)-\mathbb{E}[\hat{\psi}_n(y_n,x)]}{\hat{g}_n(x)\psi(y_n,x)}\right]$ and $\left[\frac{\hat{g}_n(x)-\mathbb{E}[\hat{g}_n(x)]}{\hat{g}_n(x)}\right]$.

Weak consistency of the kernel estimator of the point distribution function $g(.)$

If $nh_n^d \longrightarrow \infty$ as $n \longrightarrow \infty$, then, for all $x \in \mathbb{R}^d$,

1. Bias convergence

$$\mathbb{E}\left[\hat{g}_n(x)\right] - g(x) = O(h_n).$$

2. Asymptotic distribution

$$\left(nh_n^d\right)^{1/2}\left(\hat{g}_n(x) - \mathbb{E}\left[\hat{g}_n(x)\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, g(x)\|K\|_2^2\right).$$

(see Collomb (1976), Prop. 2.1 and 2.2).

Properties of $\hat{\psi}_n(y_n, x)$ (Daouia, Gardes, Girard & Lekina (2009))

If

$$y_n \to \infty \text{ such that } h_n \log y_n \to 0 \text{ and } nh_n^d \bar{F}(y_n|x) \to \infty \text{ as } n \to \infty,$$

then,

$$\text{for all } x \in \mathbb{R}^d \text{ and for all } j = 1, .., J \text{ such that } y_{n,j} = a_j y_n(1 + o(1)),$$

1. Bias convergence

$$\left\{ \mathbb{E}\left[\hat{\psi}_n(y_{n,j}, x)\right] \right\}_{\{j=1,\dots,J\}} = \{\psi(y_{n,j}, x)(1 + O(h_n \log y_n))\}_{\{j=1,\dots,J\}} .$$

2. Asymptotic distribution

$$\left\{ \sqrt{nh_n^d \psi(y_n|x)} \left(\frac{\hat{\psi}_n(y_{n,j}, x) - \mathbb{E}\left[\hat{\psi}_n(y_{n,j}, x)\right]}{\psi(y_{n,j}, x)} \right) \right\}_j \xrightarrow{\mathcal{D}} \mathcal{N}\left(0_{\mathbb{R}^J}, \|K\|_2^2 C(x)\right)$$

where $C_{j,j'}(x) = a_{j \wedge j'}^{1/\gamma(x)} \; \forall \; (j, j') \in \{1, \dots, J\}^2$ with $a_j > 0$.

Reminder ....

$$\frac{\hat{\bar{F}}_n(y_n|x)}{\bar{F}(y_n|x)} - 1 = \left[\frac{\hat{\psi}_n(y_n,x) - \mathbb{E}\left[\hat{\psi}_n(y_n,x)\right]}{\hat{g}_n(x)\psi(y_n,x)}\right] + \left[\frac{\mathbb{E}\left[\hat{\psi}_n(y_n,x)\right] - \psi(y_n,x)}{\hat{g}_n(x)\psi(y_n,x)}\right]$$
$$- \left[\frac{\hat{g}_n(x) - \mathbb{E}\left[\hat{g}_n(x)\right]}{\hat{g}_n(x)}\right] - \left[\frac{\mathbb{E}\left[\hat{g}_n(x)\right] - g(x)}{\hat{g}_n(x)}\right].$$

Asymptotic distribution of $\hat{\bar{F}}_n(y_n|x)$ (Daouia, Gardes, Girard & Lekina (2009))

If

$$y_n \to \infty \text{ such that } h_n \log y_n \to 0 \text{ and } nh_n^d \bar{F}(y_n|x) \to \infty \text{ as } n \to \infty,$$

then,

$$\text{for all } x \in \mathbb{R}^d \text{ and for all } j = 1,..,J \text{ such that } y_{n,j} = a_j y_n(1 + o(1)),$$

$$\left\{\sqrt{nh_n^d \bar{F}(y_n|x)}\left(\frac{\hat{\bar{F}}_n(y_{n,j}|x)}{\bar{F}(y_{n,j}|x)} - 1\right)\right\}_{\{j=1,\dots,J\}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0_{\mathbb{R}^J}, \frac{\|K\|_2^2}{g(x)} C(x)\right).$$

Asymptotic distribution of $\hat{q}_n(\alpha_n|x)$ (Daouia, Gardes, Girard & Lekina (2009))

If,

$$\alpha_n \to 0 \text{ and } nh_n^d \alpha_n \to \infty \text{ as } n \to \infty,$$

then,

$$\text{for all } x \in \mathbb{R}^d \text{ and for } j = 1, .., J \text{ such that } \alpha_{n,j} = \tau_j \alpha_n (1 + o(1)),$$

$$\left\{ \sqrt{nh_n^d \alpha_n} \left(\frac{\hat{q}_n(\alpha_{n,j}|x)}{q(\alpha_{n,j}|x)} - 1 \right) \right\}_{\{j=1,\ldots,J\}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0_{\mathbb{R}^J}, \gamma^2(x) \frac{\|K\|_2^2}{g(x)} \Sigma \right),$$

where $\Sigma_{j,j'}(x) = 1/\tau_{j \wedge j'}$ for $(j, j') \in \{1, \ldots, J\}^2$ with $\tau_j > 0$.

Remarks: the asymptotic variance

- is inversely proportional to $nh_n^d \alpha_n$, the estimation remains more stable when the extreme quantile is far from the boundary of the sample;
- is proportional to $\gamma^2(x) \Rightarrow$ a difficult estimation of $q(\alpha_n|x)$ for a large value of the tail index.

An application: kernel Pickands estimator

$$\hat{\gamma}_n(x) = \frac{1}{\log 2} \log \left(\frac{\hat{q}_n(k_n/n|x) - \hat{q}_n(2k_n/n|x)}{\hat{q}_n(2k_n/n|x) - \hat{q}_n(4k_n/n|x)} \right),$$

- Intermediate sequence $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$.

Asymptotic distribution of $\hat{\gamma}_n(\alpha_n|x)$ (Daouia, Gardes, Girard & Lekina (2009))

If

- The bias function $|\varepsilon(y|x)| := \left| y \frac{\ell'(y|x)}{\ell(y|x)} \right|$ is ultimately non-increasing,
- $k_n h_n^d \to \infty$ and $\sqrt{k_n h_n^d}\, \varepsilon \left(q(2k_n/n|x)|x \right) \to 0$ as $n \to \infty$,

then,

$$\sqrt{k_n h_n^d} \left(\hat{\gamma}_n(x) - \gamma(x) \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left(0, \frac{\|K\|_2^2}{g(x)} \frac{\gamma^2(x)(2^{2\gamma(x)+1}+1)^2}{4(\log 2)^2(2^{\gamma(x)}-1)^2} \right).$$

- The asymptotic variance is, up to the scale factor $\|K\|_2^2/g(x)$, the variance of the classical Pickands estimator.

Numerical experiments on simulated data

- We have generated $m = 100$ replications of the sample $\{(X_i, Y_i), i = 1, ..., n\}$ of size $n = 500$ with the conditional quantile defined by
$$q(\alpha_n|x) = (-\log \alpha_n)^{-\gamma(x)} \quad \text{(Fréchet distribution)}.$$
- We focus on the estimation of the quantile of order $\alpha_n = 5 \log (n)/n$.
- The following kernel is chosen
$$K(x) = \frac{15}{16}\left(1 - x^2\right)^2 \mathbb{1}\{|x| \leq 1\} \quad \text{(Biquadratic-kernel)}.$$

Numerical experiments on simulated data

For choosing the bandwidth h_n, we compared two strategies:

1. Yao criterion (1999),

$$h_{cv} = \arg\min_{h\in\mathcal{H}} \sum_{i=1}^{n}\sum_{j=1}^{n}\left\{\mathbb{1}\{Y_i \geq Y_j\} - \hat{\bar{F}}_{n,-i}(Y_j|X_i)\right\}^2,$$

   where $\hat{\bar{F}}_{n,-i}$ is the estimator of $\bar{F}$ computed from the sample $\{(X_k, Y_k), k = 1, \ldots, n\}$ without the *ith* observation (X_i, Y_i).

2. Oracle strategy

$$h_{oracle} = \arg\min_{h\in\mathcal{H}} \Delta\left(q(\alpha_n|.), \hat{q}_n(\alpha_n|.)\right),$$

   with

$$\Delta\left(q(\alpha_n|.), \hat{q}_n(\alpha_n|.)\right) = \left\{\frac{1}{L}\sum_{l=1}^{L}\left(q(\alpha_n|t_l) - \hat{q}_n(\alpha_n|t_l)\right)^2\right\}^{1/2}$$

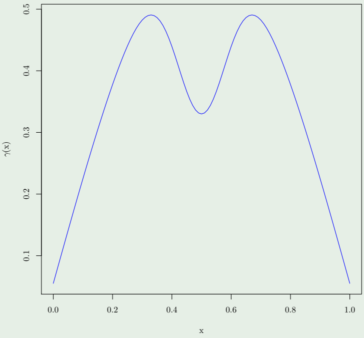
   and $t_1, \ldots, t_l$ are regularly distributed on $[0, 1]$.

Numerical experiments on simulated data

- Tail index function

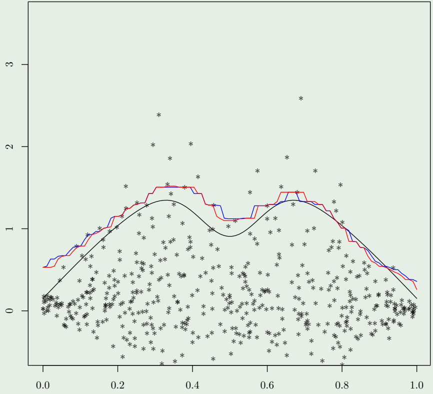
$$x \in [0,1] \mapsto \gamma(x) = \frac{1}{2}\left(\frac{1}{10} + \sin(\pi x)\right)\left(\frac{11}{10} - \frac{1}{2}\exp\left(-64\,(x-1/2)^2\right)\right).$$

Shape of function $\gamma(x)$

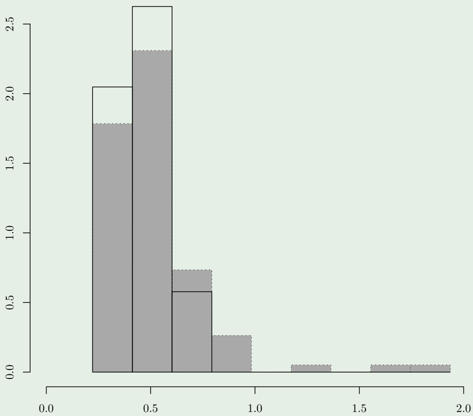


Comparison of the true quantile with the ones obtained by the Yao strategy and Oracle strategy: median estimator.

— $q(\alpha_n|x)$ — $\hat{q}_{n,oracle}(\alpha_n|x)$ — $\hat{q}_{n,cv}(\alpha_n|x)$



Comparison between the error distributions obtained with the Yao strategy (light gray) and the oracle strategy (transparent).



References

- A. Daouia, L. Gardes, S. Girard and A. Lekina. Kernel estimators of extreme level curves. http://hal.inria.fr/inria-00393588/fr/, 2009.
- A. Berlinet, A. Gannoun and E. Matzner-Løber. Asymptotic normality of convergent estimates of conditional quantiles. *Statistics*, 35:139–169, 2001.
- N.H. Bingham, C.M. Goldie and J.L. Teugels. *Regular Variation*, Cambridge University Press, 1987
- G. Collomb. *Estimation non paramétrique de la régression par la méthode du noyau*. PhD thesis, Université Paul Sabatier de Toulouse, 1976.
- J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3:119–131, 1975.
- C.J. Stone. Consistent nonparametric regression (with discussion). *The Annals of Statistics*, 5:595–645, 1977.
- W. Stute. Conditional empirical processes. *The Annals of Statistics*, 14:638–647, 1986.
- Q. Yao. Conditional predictive regions for stochastic processes. *Technical report*, University of Kent at Canterbury, 1999.