Estimation of the Angular Density in Multivariate Generalized Pareto Models

René Michel
michel@mathematik.uni-wuerzburg.de

Institute of Applied Mathematics and Statistics
University of Würzburg, Germany

18.08.2005 / EVA 2005

The Multivariate Generalized Pareto Distribution

Let $X = (X_1, \dots, X_d) \in (-\infty, 0)^d$, $d \in \mathbb{N}$, be a random vector, which has a distribution function W with the representation

$$W(x) = 1 + \left(\sum_{i=1}^{d} x_i\right) \cdot D\left(\frac{x_1}{\sum_{i=1}^{d} x_i}, \dots, \frac{x_{d-1}}{\sum_{i=1}^{d} x_i}\right)$$

close to 0. Then X follows a **generalized Pareto distribution** (GPD) with uniform margins with the Pickands dependence function $D : R_{d-1} \to [0,1]$, R_{d-1} being the unit simplex in $\mathbb{R}^{d-1}$.

The Pickands Dependence Function

D has the well known representation

$$D(t_1, \ldots, t_{d-1}) = \int_{R_{d-1}} \max\left(u_1 t_1, \ldots, u_{d-1} t_{d-1}, u_d\left(1 - \sum_{i \le d-1} t_i\right)\right) d\mu(u)$$

where $u_d = 1 - \sum_{i \le d-1} u_i$ with a measure μ on R_{d-1}.
We have $\mu(R_{d-1}) = d$.

The Angular Density

The distribution function

$$L(z_1, \dots, z_{d-1}) = \mu\left([0, z_1] \times \cdots \times [0, z_{d-1}]\right)$$

of the measure μ is called **angular distribution.** If this measure possesses a density it is called the **angular density** l. Let $(X_1, \dots, X_d)$ be a random vector following a GPD. Then

$$D(t_1, \dots, t_{d-1}) = 1 \quad \Longleftrightarrow \quad \mu(\{0\}) = 1 = \mu\{e_i\}$$

$$D(t) = \max\left(t_1, \dots, t_{d-1}, 1 - \sum_{i=1}^{d-1} t_i\right) \quad \Longleftrightarrow \quad \mu\left\{\left(\frac{1}{d}, \dots, \frac{1}{d}\right)\right\} = d$$

with e_i being the standard unit vectors in $\mathbb{R}^{d-1}$, the so called cases of **independence** and **complete dependence**.

Pickands Coordinates

Let (X_1, X_2) be a bivariate random vector with $X_1 < 0, X_2 < 0$ which is distributed by a GPD. The random variables

$$Z := X_2/(X_1 + X_2) \qquad \text{and} \qquad C := X_1 + X_2$$

are the **Pickands coordinates** of (X_1, X_2).
Z is the **angular component** and C is the **radial component**.

Conditional Density of Z

For any threshold c_0 close enough to 0 we have that the density of Z, conditional on $C > c_0$, does not depend on c_0. We denote this density by $f(z)$.

One can show that

$$g(z) := \frac{f(z)}{z(1-z)} = \text{constant } \cdot l(z),$$

$l(z)$ being the angular density.

Estimation of l

Suppose now that we have n independent copies (X_{1i}, X_{2i}) of (X_1, X_2) and denote by

$$Z_i := X_{2i}/(X_{1i} + X_{2i}), C_i := X_{1i} + X_{2i}$$

the corresponding Pickands coordinates.
Fix a threshold c_0 close to 0 and consider only those observations (X_{1i}, X_{2i}) with $C_i > c_0$. Denote these by $(\tilde{X}_{11}, \tilde{X}_{21}), \dots, (\tilde{X}_{1m}, \tilde{X}_{2m})$, where m is the random number of observations $C_i > c_0$.

Estimation of l

A natural estimator of f is, therefore, the kernel density estimator with kernel function k and bandwidth $h > 0$

$$\hat{f}_m(z) := \frac{1}{mh} \sum_{j=1}^{m} k\left(\frac{z - \tilde{Z}_j}{h}\right).$$

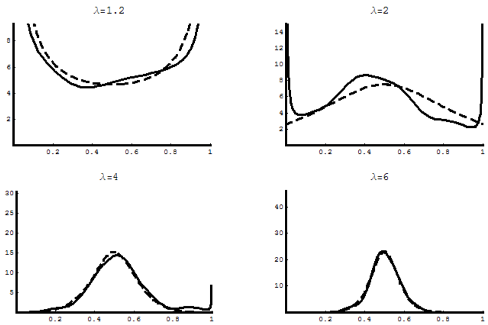
We can thus estimate a constant multiple of the angular density l by

$$\hat{g}_m(z) := \frac{\hat{f}_m(z)}{z(1-z)}.$$

Simulation Study

Applying this estimator to 50 data points generated by a simulation algorithm for the logistic distribution (Michel 2004), whereby taking *k* to be the normal kernel and using automatic bandwidth selection we get quite good results.

dashed line: underlying g, solid line: estimated g



The Case $d > 2$

The obvious generalization to $d = 3$ would be to compute threedimensional Pickands coordinates $Z_1 := X_2/(X_1 + X_2 + X_3)$, $Z_2 := X_3/(X_1 + X_2 + X_3)$ and $C := X_1 + X_2 + X_3$, take the density $f(z_1, z_2)$ of (Z_1, Z_2) and set

$$g(z_1, z_2) := \frac{f(z_1, z_2)}{z_1 z_2 (1 - z_1 - z_2)}.$$

But then we do **NOT** have

$$g(z_1, z_2) = \text{constant} \cdot l(z_1, z_2).$$

Modified Pickands Coordinates

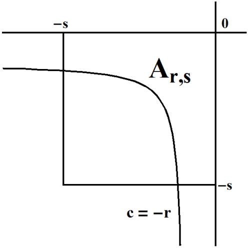
Define instead the transformation

$$T_M(x) := \left(\frac{\frac{1}{x_1}}{\frac{1}{x_1} + \ldots + \frac{1}{x_d}}, \ldots, \frac{\frac{1}{x_{d-1}}}{\frac{1}{x_1} + \ldots + \frac{1}{x_d}}, \frac{1}{x_1} + \ldots + \frac{1}{x_d} \right) =: (z_M, c_M).$$

T_M is called **transformation to modified Pickands coordinates** $z_M = (z_1, \ldots, z_{d-1}), c_M$.

The Threshold Set $A_{r,s}$

We consider those coordinates under the condition that $X \in A_{r,s}$, i.e., $||X||_\infty < s$ and $C_M < -r$.



$Q_{r,s}$ and $\chi(r,s)$

Set in addition

$$Q_{r,s} := \left\{ z \in R_{d-1} \,\middle|\, z_i > \frac{1}{rs}, i = 1, \dots, d-1, \sum_{i=1}^{d-1} z_i < 1 - \frac{1}{rs} \right\}$$

and

$$\chi(r,s) := \int_{Q_{r,s}} l(z)\, dz.$$

Then

$$Q_{r,s} \to_{r\to\infty} R_{d-1}$$

and

$$\chi(r,s) \to_{r\to\infty} d.$$

Main Result

Conditional on $X \in A_{r,s}$ the modified Pickands coordinate Z_M has the density

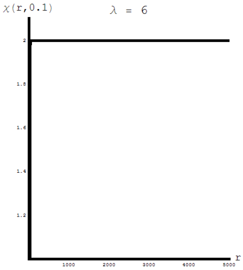
$$\frac{l(z)}{d} + O\left(d - \chi(r,s)\right) \to_{r\to\infty} \frac{l(z)}{d}.$$

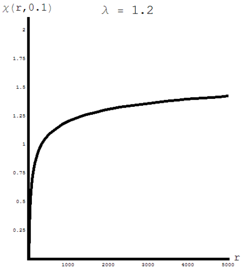
Convergence of χ

The convergence of χ to d can be very different, see for example the logistic case

$$D_\lambda(t_1,\ldots,t_{d-1}) = \left(\sum_{i=1}^{d-1} t_i^\lambda + \left(1-\sum_{i=1}^{d-1} t_i\right)^\lambda\right)^{1/\lambda}, \qquad \lambda \in [1,\infty)$$

with $\lambda = 6$ and $\lambda = 1.2$.





Estimation of l

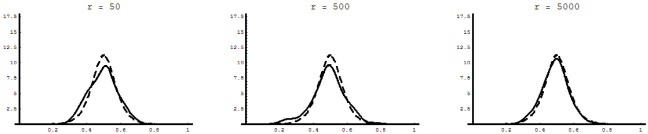
We can now estimate $l(z)$ by a multivariate kernel density estimator with data sphering

$$\hat{l}_{m,r}(z) = d \cdot \frac{1}{(\det S_m)^{1/2}\, m h^{d-1}} \sum_{i=1}^{m} k\left(\frac{S_m^{-1/2}\left(z - Z_M^{(i)}\right)}{h}\right),$$

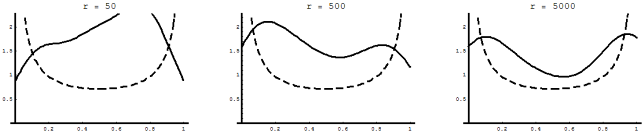
taking thereby only observations with $X_i \in A_{r,s}$ and S_m being the sample covariance matrix of the $Z_M^{(i)}$.
Under suitable regularity conditions we have asymptotic normality of $\hat{l}_{m,r}(z)$ for $m \to \infty$, $r \to \infty$.

Simulation Study

We show simulations of the estimator for various r for the logistic case with $\lambda = 6$ and $\lambda = 1.2$ in two and three dimensions. We begin with $d = 2$ and $\lambda = 6$, getting good and stable results.

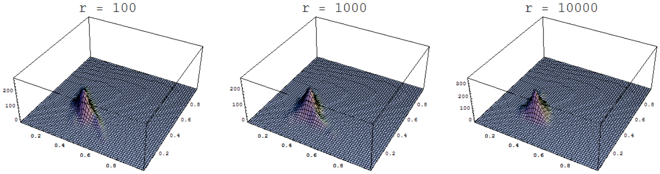


For $d = 2$ and $\lambda = 1.2$ we see varying results and slow convergence.

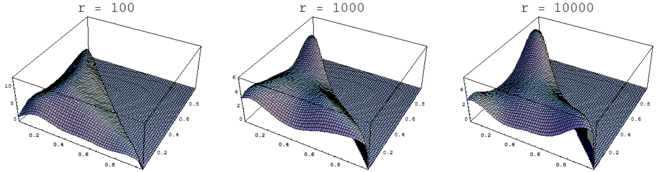


Simulation Study

For $d = 3$ and $\lambda = 6$ we get again good and stable results.



But for $d = 3$ and $\lambda = 1.2$ we see again varying results and slow convergence.



Results

If we are close to the case of complete dependence the estimator works fine, but close to the case of independence the estimator is only converging slowly to the desired result.
This estimator was also used in Coles and Tawn (1991, 1994) and Coles, Heffernan & Tawn (1999) for the extreme value case. We see here that it is also applicable in the Generalized Pareto case.
In the bivariate case a modified version of this estimator is able to deliver good results, independent of the threshold.

Thank you very much for your attention !