

Simulation Study

Let $(J_i)_{i \in \mathbb{N}}$ be a strictly stationary sequence with marginal distribution F_J and right endpoint x_R and an i.i.d. sequence of heavy-tailed random variables $(W_i)_{i \in \mathbb{N}}$ the waiting times between the observations J_i which are independent of J_i for $i \in \mathbb{N}$. Moreover, we assume that the W_i 's are in the domain of attraction of a strictly stable random variable D with laplace transform $\mathbb{E}(e^{-sD}) = \exp(-s^\beta)$, $0 < \beta < 1$, i.e., there exists a regularly varying function $b \in RV(-1/\beta)$ such that

$$b(n)(W_1 + \dots + W_n) \xrightarrow{d} D. \quad (1)$$

Given a threshold $u \in (x_L, x_R)$, consider the stopping time

$$\tau(u) := \min\{k \mid J_k > u\} \quad (2)$$

and define

$$T(u) = \sum_{k=1}^{\tau(u)} W_k \quad (3)$$

the waiting time until the sequence (J_i) exceeds the threshold u for the first time. We will consider $T(u)$ as an estimator for the inter-arrival time of exceedances of u . Under some assumptions Hees & Fried (2019) found out that

$$\frac{T(u)}{b(1/\bar{F}_J(u))} \xrightarrow{d} Z_{\beta,\theta} \quad \text{as } u \rightarrow x_R \quad (4)$$

with $\bar{F}_J := 1 - F_J$ and where $Z_{\beta,\theta}$ follows a mixture of the point measure in 0 and a Mittag-Leffler distribution, more precisely

$$Z_{\beta,\theta} \sim (1 - \theta)\varepsilon_0 + \theta\mathbb{P}_{\beta,\theta} \quad (5)$$

with $\mathbb{P}_{\beta,\theta} = ML(\beta, \theta^{-1/\beta})$. Therefore, we assume that for u high enough $T(u)$ is approximately distributed like $X_{\beta,\theta,u} := b(1/\bar{F}_J(u)) \cdot Z_{\beta,\theta} \sim (1 - \theta) \cdot \varepsilon_0 + \theta \cdot ML(\beta, b(1/\bar{F}_J(u)) \cdot \theta^{-1/\beta})$.

Since the Mittag-Leffler distribution is heavy-tailed such that there are no finite moments, we look at the fractional moments of $X_{\beta,\theta,u}$. The q -th fractional moment of X with $q \in (0, 1)$, $q < \beta$, is given by

$$m_q := m_q(X_{\beta,\theta,u}) := \mathbb{E}(X_{\beta,\theta,u}^q) = \theta \cdot \frac{q \cdot \pi \cdot (\theta^{-1/\beta} \cdot b(1/\bar{F}_J(u)))^q}{\beta \cdot \Gamma(1 - q) \sin(\frac{q\pi}{\beta})} \quad (6)$$

Hence, by calculating

$$\frac{m_q^2}{m_{2q}} = \theta \cdot \frac{\frac{q\pi}{\beta} \cdot \Gamma(1 - 2q) \sin(\frac{2q\pi}{\beta})}{2 \cdot \Gamma(1 - q)^2 \sin(\frac{q\pi}{\beta})^2}, \quad \text{for } 2q < \beta \quad (7)$$

we get rid of the unknown regularly varying function b and the survival propability $\bar{F}_J(u)$. Since $\lim_{x \rightarrow 0} \frac{\sin(x\pi)}{x\pi} = 1$ (easily shown by using L'Hopital's rule) and $\Gamma(1) = 1$ it follows that

$$\lim_{q \rightarrow 0} \frac{m_q^2}{m_{2q}} = \theta. \quad (8)$$

Equations (7) and (8) lead to two ways of estimating β and θ without the need of knowing $b \in RV(-1/\beta)$ or $\bar{F}_J(u)$. By using the empirical fractional moments

$$\hat{m}_q = \frac{1}{k_u} \sum_{i=1}^{k_u} T_i^q(u) \quad (9)$$

where $T_i(u)$, $i = 1, \dots, k_u$ are the inter-arrival times of the threshold u exceedances and k_u is the number of exceedances, we get the following estimator for θ depending on β :

$$\hat{\theta}_q(\beta) = \frac{\hat{m}_q^2}{\hat{m}_{2q}} \cdot \frac{2 \cdot \Gamma(1-q)^2 \sin(\frac{q\pi}{\beta})^2}{\frac{q\pi}{\beta} \cdot \Gamma(1-2q) \sin(\frac{2q\pi}{\beta})}. \quad (10)$$

Hence, we have to estimate β first by calculating the root of $\hat{\theta}_{q_1}(\beta) - \hat{\theta}_{q_2}(\beta)$ on $(2 \cdot \max(q_1, q_2), 1]$ (we choose $2 \cdot \max(q_1, q_2)$ as the lower interval limit because $\beta > 2 \cdot q_1, 2 \cdot q_2$ has to be fulfilled, otherwise the fractional moments don't exist).

Alternatively, we can use the result of (8) by choosing a very small fraction q close to zero 0. Then, the estimator for θ given by

$$\tilde{\theta}_q = \frac{\hat{m}_q^2}{\hat{m}_{2q}} \quad (11)$$

is independent of β . Afterwards we calculate $\hat{\beta}$ by solving (10) for β with $\tilde{\theta}_q$ on the right side of the equation.

Next to the choice of q or q_1 and q_2 the choice of the threshold u is important. We have to choose u high enough such that the limiting result of (1) holds, but the higher we choose u , the fewer observations we have for the estimates of the parameters β and θ .

The above written are the theoretical results, now we want to test them by doing a small simulation study in R:

Therefore we choose the *Max-autoregressive (MAR) Processe* with extremal index $\theta \in (0, 1)$ for $(J_i)_{i \in \mathbb{N}}$ which is defined by

$$J_1 := X_1 / \theta \quad (12)$$

$$J_i := (1 - \theta) J_{i-1} \vee X_i \quad \text{for } n \geq 2 \quad (13)$$

where $(X_i)_{i \in \mathbb{N}}$ are i.i.d. unit Fréchet distributed random variables (Ferro & Segers, 2003). The associated waiting time process $(W_i)_{i \in \mathbb{N}}$ is modeled as independent Pareto distributed random variables with c.d.f.

$$F_W(x) := 1 - \left(\frac{s}{x}\right)^\beta \quad \text{for } 0 < \beta < 1, s > 0 \text{ and } x \geq s. \quad (14)$$

The Pareto distribution with $s = \Gamma(1 - \beta)^{-1/\beta}$ is in the domain of attraction of D such that $b(n) = n^{-1/\beta}$ holds in (1) (Meerschaert & Sikorskii, 2012).