Risk Analytics

Extreme Value Theory
Block Maxima Approach

Frédéric Aviolat & Juraj Bodik

2025

Say you have observed outcomes $X_1, \ldots, X_n$ where higher values are worse, e.g.:

- Extreme temperatures (heatwaves, droughts, wild fires)
- Storms, extreme winds (hail)
- Heavy rainfalls, extreme discharges (floods)

Extreme value theory:

- How bad does it get?
- When should we ring the alarm? (see Challenger example)

How bad does it get?

"Study the worst case"

Two classical approaches:

- Block maxima: $\max(X_1, \ldots, X_n)$
- Peaks-over-Threshold: $\{X : X > u\}$ (for a high threshold u)

Study of Maxima

Let $X_1, X_2, \ldots$ be iid random variables with distribution function (df) F,

$$F(x) = \mathbb{P}(X \leq x).$$

Let $M_n = \max(X_1, \ldots, X_n)$ be worst–case value among $1, \ldots, n$.

$$\begin{aligned}
\mathbb{P}(M_n \leq x) &= \mathbb{P}(X_1 \leq x, \ldots, X_n \leq x) \\
&= \mathbb{P}(X_1 \leq x)\, \mathbb{P}(X_2 \leq x) \cdots \mathbb{P}(X_n \leq x) \\
&= F^n(x)\,.
\end{aligned}$$

It can be shown that, almost surely, $M_n \overset{n\to\infty}{\longrightarrow} x_F$, where $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\} \leq \infty$ is the right endpoint of F.

$\Rightarrow$ But what about normalized maxima?

Limiting Behaviour of Sums or Averages

We are familiar with the central limit theorem:

Let $X_1, X_2, \ldots$ be iid with finite mean μ and finite variance σ^2. Let $S_n = X_1 + X_2 + \ldots + X_n$. Then

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq x\right) \xrightarrow{n\to\infty} \Phi(x)\,,$$

where Φ is the distribution function of the standard normal distribution

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du\,.$$

The distribution of the normalized sum of the X_i converges to the standard normal distribution.

Limiting Behaviour of Sample Extrema

Let $X_1, X_2, \ldots$ be iid from F and let $M_n = \max(X_1, \ldots, X_n)$.

Suppose we can find sequences of real numbers $a_n > 0$ and b_n such that $(M_n - b_n)/a_n$, the sequence of normalized maxima, converges in distribution, i.e.

$$\begin{aligned} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) &= \mathbb{P}(M_n \leq a_n x + b_n) \\ &= F^n(a_n x + b_n) \\ &\overset{n\to\infty}{\longrightarrow} H(x), \end{aligned}$$

for some non-degenerate distribution function $H(x)$.

We say that F is in the maximum domain of attraction of H, or $F \in \mathrm{MDA}(H)$.

In fact, we can find the form of $H(x)$ (Fisher-Tippett theorem).

Generalized Extreme Value Distribution

The GEV has distribution function

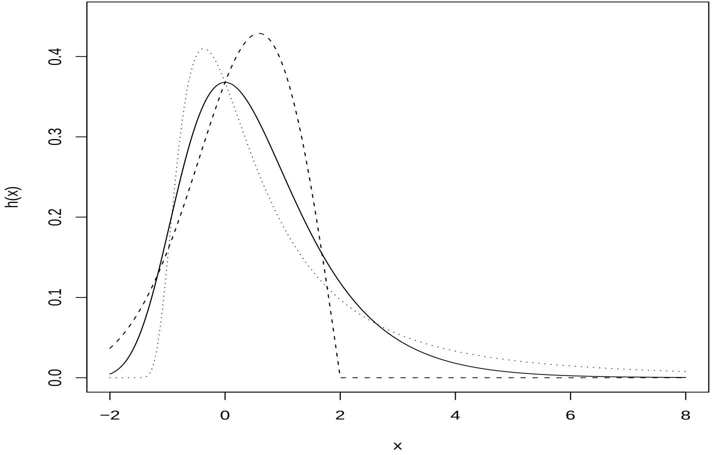
$$H_{\mu,\sigma,\xi}(x) = \begin{cases} \exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right) & \text{if } \xi \neq 0\,, \\ \exp\left(-e^{-(x-\mu)/\sigma}\right) & \text{if } \xi = 0, \end{cases}$$

where $1+\xi(x-\mu)/\sigma > 0$.

- μ is the location parameter.
- σ is the scale parameter.
- ξ is the shape parameter.

$$\text{For } \left\{ \begin{array}{c} \xi > 0 \\ \xi = 0 \\ \xi < 0 \end{array} \right\} \text{ we also say that H is } \left\{ \begin{array}{c} \text{Fréchet} \\ \text{Gumbel} \\ \text{Weibull} \end{array} \right\}.$$

GEV Densities



Solid line corresponds to $\xi = 0$ (Gumbel); dotted line is $\xi = 0.5$ (Fréchet); dashed line is $\xi = -0.5$ (Weibull). $\mu = 0$ and $\sigma = 1$.

Fisher–Tippett Theorem (1928)

Theorem If $F \in \mathrm{MDA}(H)$ then H is of the type $H_{\mu,\sigma,\xi}$ for some μ, σ, ξ.

"If suitably normalized maxima converge in distribution to a non–degenerate limit, then the limit distribution must be an extreme value distribution."

Remark Essentially all commonly encountered continuous distributions are in the maximum domain of attraction of an extreme value distribution.

Fisher-Tippett: Examples

Recall: $F \in \mathrm{MDA}(H_\xi)$, iff there are sequences a_n and b_n with

$$\mathbb{P}\left\{(M_n - b_n)/a_n \le x\right\} = F^n(a_n x + b_n) \stackrel{n\to\infty}{\longrightarrow} H_\xi(x)\,.$$

We have the following examples:

- The exponential distribution,

  $$F(x) = 1 - e^{-\lambda x}, \lambda > 0, x \ge 0$$

  is in $\mathrm{MDA}(H_0)$ (Gumbel). Take $a_n = 1/\lambda$, $b_n = (\log n)/\lambda$.
- The Pareto distribution,

  $$F(x) = 1 - \left(\frac{\kappa}{\kappa + x}\right)^\alpha, \quad \alpha, \kappa > 0, \quad x \ge 0,$$

  is in $\mathrm{MDA}(H_{1/\alpha})$ (Fréchet). Take $a_n = \kappa n^{1/\alpha}/\alpha$, $b_n = \kappa n^{1/\alpha} - \kappa$.

Using Fisher–Tippett on data: Block Maxima Method

If you are given n values, use the limiting distribution to model the max M_n: first find the a_n and b_n from the theorem, and then

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) \approx H_{0,1,\xi}(x)\,,$$

or $\mathbb{P}(M_n \leq y) = H_{b_n,a_n,\xi}(y)$.

+ All that's left is to estimate three parameters: ξ, b_n and a_n
− Need repeated values of $M_n \Rightarrow$ required data is a multiple of n

Remark

The values b_n and a_n are equivalent to the parameters μ and σ in the formula.

ML Inference for Maxima

We have block maxima data $\mathbf{y} = \left(M_n^{(1)}, \ldots, M_n^{(m)}\right)$ from m blocks of size n. We wish to estimate $\boldsymbol{\theta} = (\xi, \mu, \sigma)$. We construct a log–likelihood by assuming we have independent observations from a GEV with density $h_{\boldsymbol{\theta}}$,

$$l(\boldsymbol{\theta}; \mathbf{y}) = \log \left\{ \prod_{i=1}^{m} h_{\boldsymbol{\theta}} \left(M_n^{(i)}\right) \mathbb{1}_{\left\{1+\xi\left(M_n^{(i)}-\mu\right)/\sigma>0\right\}} \right\},$$

and maximize this w.r.t. $\boldsymbol{\theta}$ to obtain the MLE $\widehat{\boldsymbol{\theta}} = (\widehat{\xi}, \widehat{\mu}, \widehat{\sigma})$.

In defining blocks, bias and variance must be traded off:

- we reduce bias by increasing the block size n;
- we reduce variance by increasing the number of blocks m.

Return level and return period

Return level

Let X be a random variable with df F. The return level $R_\alpha = F^{-1}(\alpha)$ is the α-quantile, for probability α.

Return level of empirical data

Let $X_1 <= \ldots <= X_n$ be a series of ordered observations with df F. The return level $R_k = F^{-1}(1 - \frac{1}{k})$ is the value exceeded in $1/k$ observations, on average.

Return period

Let X be a random variable with df F. The return period $P(x) = \frac{1}{1-F(x)} = \frac{1}{\bar{F}(x)}$ is the expected number of observations to get a value greater than x.

Return period (of empirical data)

Let $X_1 <= \ldots <= X_n$ be a series of ordered observations with df F. One can check that $P(R_k) = k, \forall k$ and $R_{P(x)} = x, \forall x$.

Return level for GEV

- The return level for probability α of $H_{\mu,\sigma,\xi}$ is

$$R_\alpha = H_{\mu,\sigma,\xi}^{-1}(\alpha) = \begin{cases} \mu - \frac{\sigma}{\xi}\left(1 - (-\log(\alpha))^{-\xi}\right) & \text{if } \xi \neq 0 \\ \mu - \sigma \log(-\log(\alpha)) & \text{if } \xi = 0 \end{cases}$$

- For a return period $T = 1/p$, take $\alpha = 1 - p$. The return level is the value expected to be exceeded on average once every T periods.
- More precisely, R_{1-p} is exceeded by the maximum in any particular period with probability p.
- For example, if we have annual maxima, R_{1-p} is the value expected to be exceeded on average once every T years, and it is exceeded by the annual maximum in any particular year with probability p.

Return level plot

- If we set $y_p = -\log(1-p)$,

$$R(y_p) = \begin{cases} \mu - \frac{\sigma}{\xi}\left(1 - y_p^{-\xi}\right) & \text{if } \xi \neq 0 \\ \mu - \sigma \log(y_p) & \text{if } \xi = 0 \end{cases}$$

  The plot $\{(\log(y_p), R(y_p)) : 0 < p < 1\}$ is the return level plot of $H_{\mu,\sigma,\xi}$.

- **Remarks**
  - If $\xi = 0$, the plot is linear.
  - If $\xi < 0$, the plot is concave with $\lim_{p\to 0} R(y_p) = \mu - \sigma/\xi$.
  - If $\xi > 0$, the plot is convex and has no finite bound.

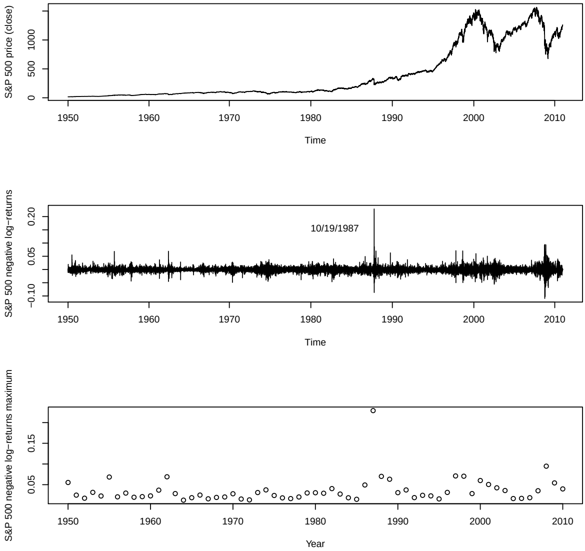
An Example: S&P 500

At our disposal are all daily closing values of the index since 1950. We analyse annual maxima of daily negative log-returns in the index. These values $M_{260}^{(1)}, \ldots, M_{260}^{(61)}$ are assumed to be iid from $H_{\xi,\mu,\sigma}$.

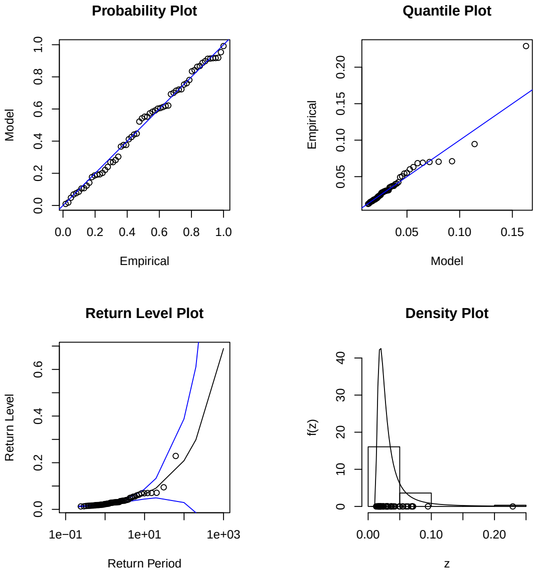
Remark

Although we have only justified this choice of limiting distribution for maxima of iid data, it turns out that the GEV is also the correct limit for maxima of stationary time series, under some technical conditions on the nature of the dependence. These conditions are fulfilled, for example, by GARCH processes.

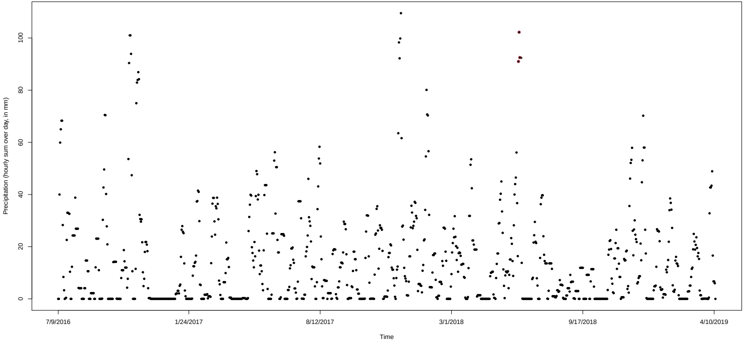
S&P 500


S&P 500 price (close)
0
500
1000
1950
1960
1970
1980
1990
2000
2010
Time
S&P 500 negative log−returns
−0.10
0.05
0.20
10/19/1987
1950
1960
1970
1980
1990
2000
2010
Time
S&P 500 negative log−returns maximum
0.05
0.15
1950
1960
1970
1980
1990
2000
2010
Year

S&P 500 diagnostic graphs



Heavy rainfall in Lausanne



Heavy rainfall in Lausanne