

Risk Analytics

Extreme value theory: summary and practice

Frédéric Aviolat & Juraj Bodik

2025

Block maxima

Let X_1, \dots, X_n be a series of iid data from F and let $M_n = \max(X_1, \dots, X_n)$.

If $F \in \text{MDA}(H)$, then:

- H is a Generalized Extreme Value Distribution (GEV), with parameters μ, σ , and ξ .
- The maxima $M_n \xrightarrow{n \rightarrow \infty} H_{\mu, \sigma, \xi}$, for some μ, σ, ξ .

Peaks-over-threshold

The POT (peaks-over-threshold) model is a limit model for threshold exceedances in iid processes. The limit is derived by considering datasets X_1, \dots, X_n and thresholds u_n that increase with n and letting $n \rightarrow \infty$. The limit model says that:

- Exceedances occur according to a homogenous Poisson process in time.
- Excess amounts above the threshold are iid and independent of exceedance times.
- Distribution of excesses is Generalised Pareto (GPD).

Stationary series

Let X_1, \dots, X_n be a **stationary** time series.

- **Block maxima** The maxima can usually be considered independant. So asymptotic results are still valid.
- **Peaks-over-threshold**
 - Asymptotic results are still valid.
 - But clustering usually happens.
 - Use declustering.
 - Take into account the **extremal index** θ .
 - Clusters have mean size $1/\theta$.

Non-stationary series

Let X_1, \dots, X_n be a **non-stationary** time series.

- **Block maxima** Use time-varying parameters:

$$\mu(t) = g(t, \alpha)$$

$$\sigma(t) = h(t, \alpha)$$

$$\xi(t) = q(t, \alpha)$$

- **Peaks-over-threshold** Use time-varying parameters:

$$u(t) = v(t, \alpha)$$

$$\beta(t) = b(t, \alpha)$$

$$\xi(t) = q(t, \alpha)$$

where $\alpha = (\alpha_0, \dots, \alpha_K)$ is a vector of parameters.

Return level and return period

Let X be a random variable with df F .

Return period

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 level

The return level $R_\alpha = F^{-1}(\alpha)$ is the α -quantile, for probability α .

The k -period return level $R_k = F^{-1}(1 - \frac{1}{k})$ is the $(1 - \frac{1}{k})$ -quantile.

Value-at-Risk and Expected Shortfall

Let X be a random variable with df F .

Value-at-Risk

$\text{VaR}_\alpha(X) = F^{-1}(\alpha)$ is the α -quantile, for probability α .

Expected Shortfall

$\text{ES}_\alpha(X) = \mathbb{E}(X \mid X > \text{VaR}_\alpha(X))$ is the average value of X , when X is higher than $\text{VaR}_\alpha(X)$.

Diagnostic plots

Let X_1, \dots, X_n be a series of observed data. Let E be the empirical distribution of this data, and \hat{F} the estimated fitted distribution.

Probability plot $\{(E(X_i), \hat{F}(X_i)) \mid i = 1, \dots, n\},$

$$E(X_i) = \frac{i}{n+1}$$

Quantile plot $\{(E^{-1}(p), \hat{F}^{-1}(p)) \mid 0 < p < 1\},$

$$E^{-1}(p) = X_{(i)}$$

Return level plot $\{(\log(y_p), R(y_p)) \mid 0 < p < 1\},$

where $y_p = -\log(1 - p)$ and $R(y_p) = \hat{F}^{-1}(y_p).$

Packages in R

The following packages allow for extreme values analysis:

- ismev
- extRemes
- POT
- evd

See the documentation of these packages and their functions.