Extreme Value Theory

Outline

1. Generalized Pareto Distribution (GPD)
2. Excess Distribution Over Threshold
3. Estimating excess distribution and applications

Overview of Extreme Value Theory

- Extreme value theory (EVT) investigates the properties of the tail distributions of random variables
- For right tail analysis of variable x, we use x directly; for left tail, we consider $-x$
- Key steps in EVT:
  1. Choose a threshold level u in the tail of the distribution
  2. Apply key theorem: For a wide class of distributions, as u increases, the conditional distribution of excesses $v - u$ given $v > u$ converges to a Generalized Pareto Distribution (GPD)
- This provides a theoretical foundation for modeling extreme events

Generalized Pareto Distribution (GPD)

- The CDF of the GPD is given by:

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

  where ξ is the shape parameter and $\beta > 0$ is the scale parameter.

- (Q) Show that, for any $x \geq 0$, $G_{\xi,\beta}(x) \to G_{0,\beta}(x)$ as $\xi \to 0^+$.

Domain of the GPD

- The support (domain) of the GPD depends on the shape parameter ξ:
  - When $\xi \geq 0$: $x \in [0, \infty)$ (unbounded upper tail)
  - When $\xi < 0$: $x \in [0, -\beta/\xi]$ (bounded upper tail)
- The scale parameter $\beta > 0$ stretches or compresses the distribution

Density and Moments of the GPD

- The density function of the GPD is given by:

$$g_{\xi,\beta}(x) = \begin{cases} \frac{1}{\beta}\left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}-1} & \text{if } \xi \neq 0, \\ \frac{1}{\beta} e^{-\frac{x}{\beta}} & \text{if } \xi = 0, \end{cases}$$

- The shape parameter ξ determines the tail behavior:
  - $\xi > 0$: Heavy-tailed distribution (Pareto-type)
  - $\xi = 0$: Exponential-type tail
  - $\xi < 0$: Light-tailed distribution with finite upper endpoint

Mean and Variance of the GPD

- The mean of the GPD exists if $\xi < 1$:

$$E[X] = \begin{cases} \frac{\beta}{1-\xi} & \text{if } \xi < 1, \\ \text{undefined} & \text{if } \xi \geq 1. \end{cases}$$

- The mean diverges for $\xi \geq 1$, indicating heavy tails
- The variance exists if $\xi < 0.5$:

$$\mathrm{Var}(X) = \begin{cases} \frac{\beta^2}{(1-\xi)^2(1-2\xi)} & \text{if } \xi < 0.5, \\ \text{undefined} & \text{if } \xi \geq 0.5. \end{cases}$$

- The variance diverges for $\xi \geq 0.5$, indicating heavy tails

Excess distribution over threshold u

- Let X be a random variable with CDF $F_X(x)$. The excess distribution over a threshold u is defined as:

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x+u) - F(u)}{1 - F(u)}, x \geq 0.$$

Stability of GPD Excess Distribution

- If $X \sim G_{\xi,\beta}$, then the excess distribution over threshold u is

$$F_u(x) = G_{\xi,\beta+\xi u}(x).$$

- This means that the excess distribution is also a GPD with modified scale parameter $\beta + \xi u$.
- The shape parameter ξ remains unchanged.

Pickands–Balkema–de Haan Theorem

- The Pickands–Balkema–de Haan theorem states that for a wide class of distributions, the excess distribution converges to a GPD as the threshold u increases, i.e., For large u, there exists ξ and β such that

$$F_u(x) \approx G_{\xi,\beta}(x), \forall x \geq 0.$$

- This means that for large enough u, the distribution of excesses behaves like a GPD.
- The parameters ξ and β can be estimated from data using methods like maximum likelihood estimation (MLE).

Estimating GPD Parameters with MLE

Given a sample data $X_1, X_2, \ldots, X_n$ from a distribution X, we can estimate ξ and β for $F_u(x) \approx G_{\xi,\beta}(x)$:

- We first choose a value for u. (A value close to the 95th percentile point of the empirical distribution usually works well.)
- Rank the data above the threshold u, reindexing them as $X_1, \ldots, X_{n_u}$, where n_u is the number of observations above u.
- The likelihood function for the GPD is given by:

$$L(\xi, \beta) = \prod_{i=1}^{n_u} g_{\xi,\beta}(X_i - u).$$

- We maximize log-likelihood:

$$\log L(\xi, \beta) = \sum_{i=1}^{n_u} \frac{1}{\beta} \left(1 + \frac{\xi(X_i - u)}{\beta} \right)^{-\frac{1}{\xi} - 1}$$

- These equations can be solved numerically using optimization techniques.

Estimating the tail probability

- Let $F_u(x) \approx G_{\xi,\beta}(x)$ be the estimated excess distribution.
- The tail probability is:

$$\begin{aligned}\mathbb{P}(X > x) &= \mathbb{P}(X - u > x - u | X > u)\mathbb{P}(X > u) \\ &= (1 - F_u(x - u)) \cdot (1 - F(u)).\end{aligned}$$

- We estimate $F_u \approx G_{\xi,\beta}$ and $1 - F(u) \approx \frac{n_u}{n}$.
- Thus, the estimated tail probability is:

$$\hat{P}(X > x) = \frac{n_u}{n}\left(1 + \xi\frac{x - u}{\beta}\right)^{-\frac{1}{\xi}}.$$

Equivalence to the power law

- The estimated tail probability is:

$$\hat{P}(X > x) = \frac{n_u}{n}\left(1 + \xi\frac{x-u}{\beta}\right)^{-\frac{1}{\xi}}.$$

- Set $u = \beta/\xi$ to simplify:

$$\hat{P}(X > x) = \frac{n_u}{n}\left(\frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}}.$$

- This is equivalent to a power law distribution with exponent $\alpha = 1/\xi$:

$$\hat{P}(X > x) \sim Cx^{-\alpha}, \text{ where } C = \frac{n_u \xi^{-\alpha}}{n\beta^{-\alpha}}.$$

- This shows that the GPD tail behaves like a power law for large x.

Calculating Value-at-Risk (VaR) using GPD

- To calculate the Value-at-Risk (VaR) at level q, we need to find the quantile v_q such that:

$$\hat{P}(X > v_q) = 1 - q.$$

- Using the estimated tail probability:

$$\frac{n_u}{n}\left(1 + \xi \frac{v_q - u}{\beta}\right)^{-\frac{1}{\xi}} = 1 - q.$$

- Solving for v_q yields:

$$v_q = u + \frac{\beta}{\xi}\left(\left(\frac{n(1-q)}{n_u}\right)^{-\xi} - 1\right).$$

Calculating Expected Shortfall (ES) using GPD

- The Expected Shortfall (ES) at level q is defined as:

$$\mathrm{ES}_q(X) = \frac{1}{1-q}\int_q^1 v_p \, dp.$$

- If $\xi \neq 1$, using the expression for v_p, we have:

$$\mathrm{ES}_q(X) = \frac{v_q + \beta - \xi u}{1-\xi}.$$