A Century of Extreme Value Theory

Roberto Vásquez Martínez [1]
Supervised by: Dr Jennifer Wadsworth [1]
[1]Lancaster University

excellence with impact

Engineering and Physical Sciences Research Council

Motivation

The study of **extreme value theory** provides a rigorous statistical framework for modelling rare but potentially catastrophic events, such as financial crashes, natural disasters, and infrastructure failures. In particular, **multivariate** extreme value theory extends this analysis to systems with multiple **interdependent variables**.

1. Univariate Extremes

The classical setting is to consider an i.i.d. sample $\{X_i, i = 1, \dots, n\}$. The objective in extreme value theory is to study the distribution of $M_n := \{X_1, \dots, X_n\}$. The main idea is to **approximate** the distribution of M_n for n large.

Fisher-Tippett-Gnedenko *(1928,1943)*

Let $\{X_n, n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables with common df F_X. Assume there exist constants $\{a_n > 0, n \in \mathbb{N}\}$ and $\{b_n, n \in \mathbb{N}\}$ such that

$$\mathbb{P}\left[\frac{M_n - b_n}{a_n} \leqslant z\right] \to G(z) \text{ non-degenerate} \quad \text{as} \quad n \to \infty,$$

Then, G belongs to the generalised extreme value (GEV) family of distributions determined by the subfamilies **Weibull**, **Gumbel**, and **Fréchet**, and we say F_X belongs to the **domain of attraction** of G

Pickands-Balkema-De Haan *(1970's)*

F_X with ω_{F_X} upper-end point belongs to a domain of attraction if and only if

$$\lim_{u \to \omega_{F_X}} \sup_{0 < z < \omega_{F_X} - u} \left| \frac{\overline{F}_X(z+u)}{\overline{F}_X(u)} - \overline{P}_{\xi, \sigma(u), \mu}(z) \right| = 0,$$

where $P_{\xi,\sigma(u),\mu}$ is the **generalised Pareto distribution**.

2. Multivariate Extremes

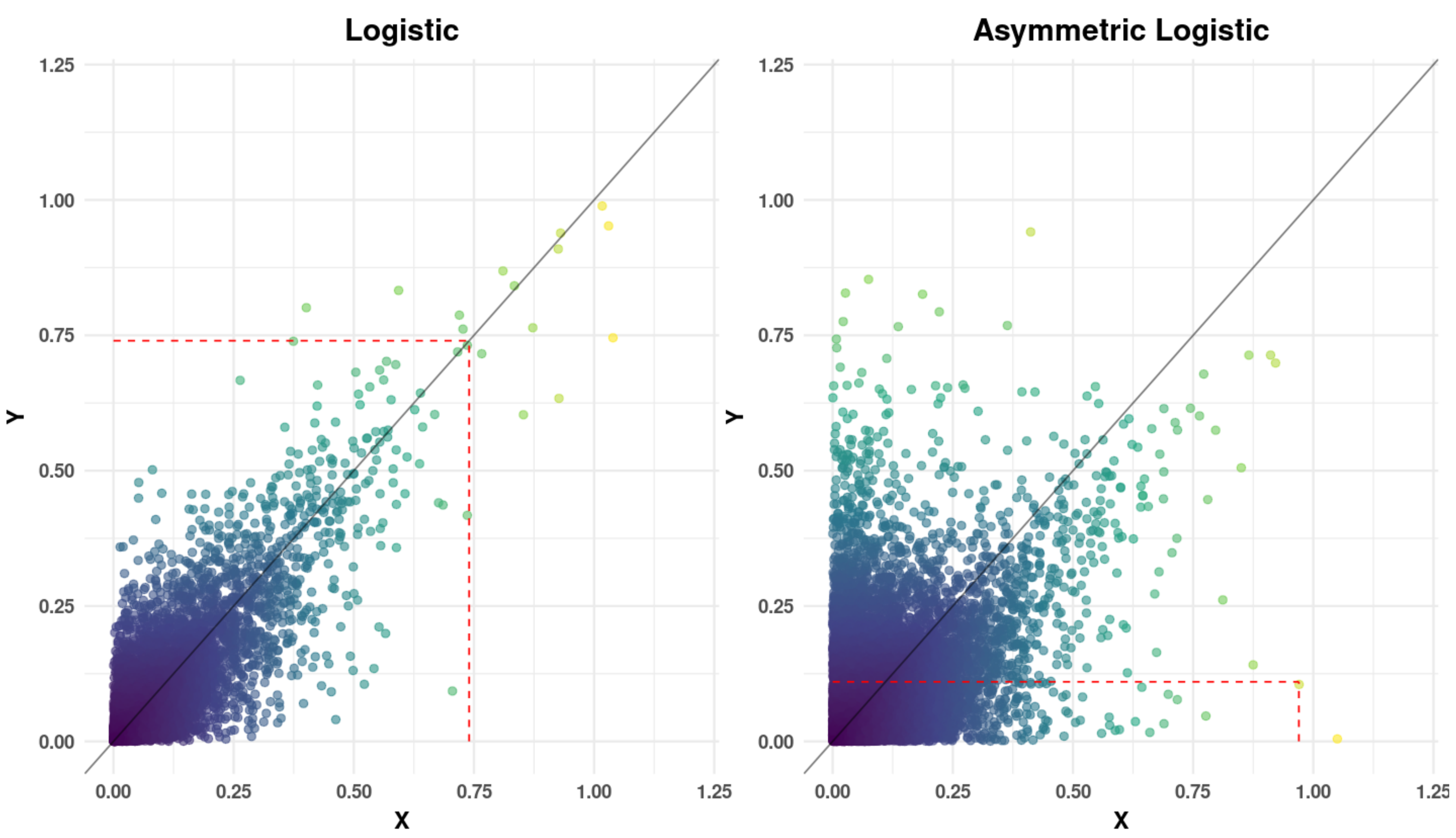
The two main considerations in the multivariate case are the **lack of ordering**, and the **dependence structure**.

Componentwise Maxima *(1980's)*

Consider an i.i.d. sample $\{\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(d)}), i = 1, \dots, n\}$ from the random vector $\mathbf{X} = (X_1, \dots, X_d)$ with distribution $F_{\mathbf{X}}$. Define the **componentwise maxima** variables $M_{k,n} := \max_{1 \leqslant i \leqslant n} X_i^{(k)}$ for $k = 1, \dots, d$, and let $\mathbf{M}_n := (M_{1,n}, \dots, M_{d,n})$. Then, the natural generalisation of the Fisher-Tippett-Gnedenko theorem considers multivariate extreme value distributions as limits of the form

$$\mathbb{P}\left[\frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n} \leqslant \mathbf{z}\right] \longrightarrow G(\mathbf{z}) \quad \text{as} \quad n \to \infty,$$

Figure 1. Examples of multivariate extreme value distributions.



Point Process Approach *(Coles and Tawn 1991)*

The idea is to consider a normalised sample cloud N_n such that $N_n \xrightarrow{d} N$ where N is a non-homogeneous Poisson process related to the distribution G. Assume $\mathbf{X}$ with support in $\mathbb{R}_+^d$ with standard Fréchet margins. First, consider **pseudo-polar coordinates**: $R := \sum_i^d X_i$ is the radial component, and $\mathbf{W} := \mathbf{X}/R$ is the angular component Then, N has intensity measure μ satisfying:

$$\mu(dr \times d\mathbf{w}) = \frac{dr}{r^2} dH(\mathbf{w}),$$

where H is a positive measure in the $(d-1)$-simplex $S_d := \{\mathbf{w} = (w_1, \dots, w_d) : \sum_{j=1}^d w_j = 1, \; w_j \geqslant 0\}$ that codes the dependence structure, and

$$G(\mathbf{z}) = \exp(-\mu(A)) \quad \text{where} \quad A = \mathbb{R}_+^d \backslash \{(0, z_1) \times \cdots \times (0, z_d)\}.$$

3. Modelling Dependence

The multivariate case's main problem is understanding the **extreme dependence** relationships between the variables studied.

Dependence measures

Let (X, Y) be a random vector with copula function C. The coefficient of extreme dependence between X and Y is defined by

$$\chi := \lim_{u \to 1^-} \mathbb{P}\left(Y > F_Y^{\leftarrow}(u) \mid X > F_X^{\leftarrow}(u)\right) = \lim_{u \to 1^-} \frac{1 - 2u + C(u,u)}{1-u}.$$

Additionally, $\bar{\chi} = \lim_{u \to 1^-} \frac{2\log(1-u)}{\log \bar{C}(u,u)} - 1$

$(\chi > 0, \bar{\chi} = 1) \leftrightarrow$ asymptotic dependence
$(\chi = 0, \bar{\chi} > 0) \leftrightarrow$ asymptotic independence

Many statistical approaches developed assume **asymptotic dependence** and struggle near the **asymptotic independence** cases.

Some alternatives:

- **Residual dependence coefficient** η (Ledford and Tawn 1996): Assuming same Fréchet margins and joint **regular variation**:

$$\mathbb{P}[X > r, Y > r] \approx L(r) r^{-\frac{1}{\eta}} \quad \text{and} \quad \bar{\chi} = 2\eta - 1.$$

  Measure of the strength of the dependence decay for the asymptotic independent variables.
- **Conditional Extremes** (Heffernan and Tawn 2004): Capture both cases. In Exponential margins, it assumes:

$$\left(\frac{Y - a_1(X)}{b_1(X)}, X - u\right) \Big| X > u \xrightarrow{d} (Z, E),$$

  where E is standard exponential independent of Z. a_1 and b_1 concentrate the dependence structure.

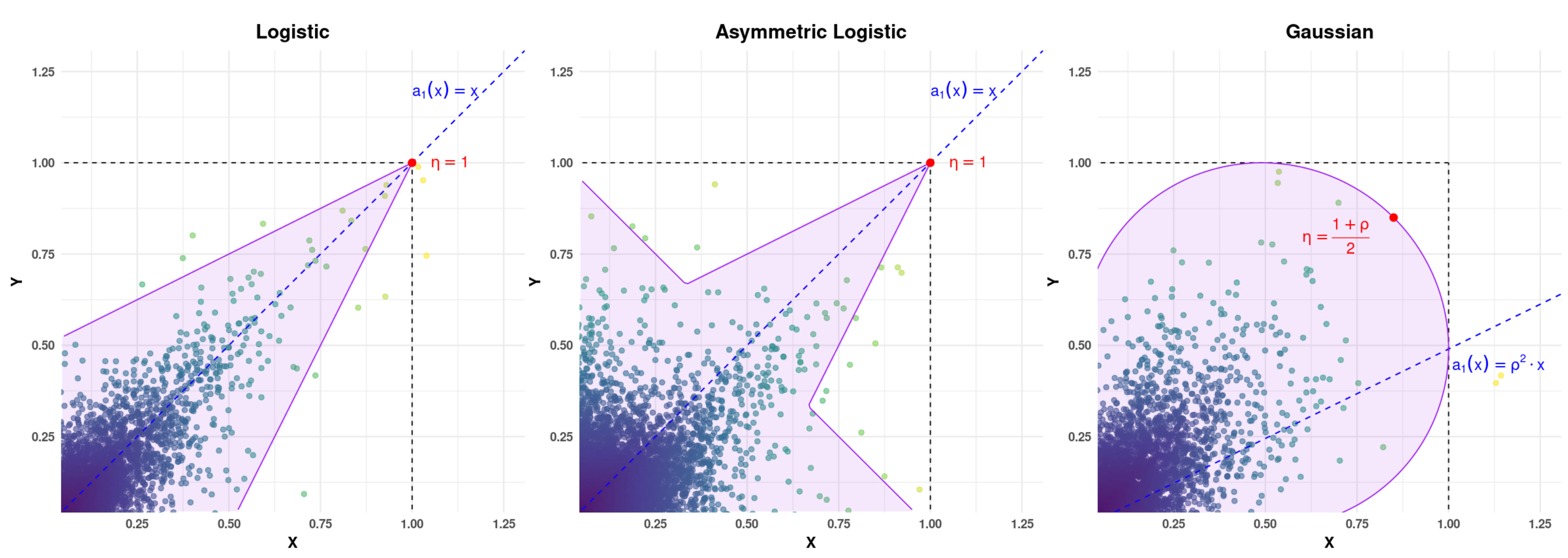
4. Geometric Approach

Provide an **unification** of the main approaches developed. Moreover,

- Embraces more general forms of extreme dependence structures, e.g., beyond regular variation.
- It is a self-consistent framework, e.g., solves the contradiction between approaches.

The idea is to provide a **complete** understanding of the join tail through a **limit set** of an appropriate normalised sample cloud $N_n = \{a_n^{-1}\mathbf{X}_i, i = 1, \dots, n\}$.

Figure 2. Limit borders (purple) and link with the other approaches



Main Result *(Nolde and Wadsworth 2022)*

Let $\mathbf{X}$ be a random vector with support in $\mathbb{R}_+^d$ and marginal distributions asymptotically equal to a von Mises function (Gumbel domain), i.e., the tail is approximately $e^{-\psi(x)}$. Assume the joint probability density $f_{\mathbf{X}}$ of $\mathbf{X}$ satisfies

$$-\frac{\log f_{\mathbf{X}}(t\mathbf{x}_t)}{\psi(t)} \longrightarrow g(\mathbf{x}), \quad \text{as} \quad t \to \infty \quad \text{and} \quad \mathbf{x}_t \longrightarrow \mathbf{x}.$$

The random set N_n converges to the limit set $G := \{\mathbf{x} \in \mathbb{R}_+^d : g(\mathbf{x}) \leqslant 1\}$, and the scaling sequence $\{a_n > 0, n \in \mathbb{N}\}$ can be chosen such that $\psi(a_n) \sim \log n$.

The function g is called **gauge function** and the level set $\partial G := \{\mathbf{x} \in \mathbb{R}_+^d : g(\mathbf{x}) = 1\}$ codes the complete joint behaviour of $\mathbf{X}$.

Further directions

- Relaxation of the i.i.d. assumption.
- Development of statistical methodology using the geometric approach.

Scan for more information...

 v.roberto@lancaster.ac.uk