Annual Progress Review

Thomas William Boughen

Newcastle University

School of Mathematics, Statistics and Physics

1 Intuition

2 Extremes

Since the aim is to gain understanding about the behaviour of the degree distribution of networks at the right tail, it seems natural to look to using methods from extreme value theory. However, networks by their nature are discrete and so it may not be best to be using methods that are usually used in relation to continuous random variables. For this reason, this section starts with a review of what theory exists for modelling the extreme values of continuous random variables before moving to details what can be used when instead considering discrete random variables as is the case for the degree distributions of random networks.

2.1 Continuous Extremes

Studying the properties of the right tail of the distribution of a continuous random variable, means that the focus is on the largest values that the random variable can take. So, a natural place to start is to consider the distribution of the block maxima of such a random variable. That is, for a set of iid random variables $\{X_1, \dots, X_n\}$ with common cumulative density function (cdf) F what is the distribution of $M_n = \max\{X_1, \dots, X_n\}$? This question is answered by the Fisher–Tippett–Gnedenko theorem REF.

Theorem 2.1.1 (Extreme Value Theorem)**.** *Let $X_1, \dots, X_n$ be a sample of iid random variables with common cdf F with block maxima $M_n = \max\{X_1, \dots, X_n\}$ and suppose that there exists $a_n > 0, b_n \in \mathbb{R}$ such that $\lim_{n\to\infty} \Pr(\frac{1}{a_n}[M_n - b_n]) = G(x)$, then F is said to be in the domain of attraction of G, denoted $F \in \mathcal{D}(G)$,and G is of one of three types:*

- *Gumbel:* $\Lambda(x) = \exp\{-\exp(-x)\}, \quad x \in \mathbb{R}$
- *Fréchet:* $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}, \quad x \geq 0, \alpha > 0$
- *Weibull:* $\Psi_\alpha(x) = \exp\{-x^{-a}\}, \quad x < 0, \alpha > 0$

While this is a useful result, it may prove difficult to find the sequences a_n, b_n in practice, so a simpler method to establish what domain of attraction a distribution belongs to would be nice. Luckily, this can be done through the concept of regular variation and is what will be used to define the domains of attraction and tail-heaviness through the rest of this report.

Definition 2.1.1 (Domains of Attraction)**.** The distribution F belongs to the Fréchet domain of attraction $\mathcal{D}(\Phi_\alpha)$ if and only if its complement (the survival function) $\bar{F}$ is regularly varying with index $-\alpha$ i.e.:

$$\bar{F}(x) = x^{-\alpha} L(x), \qquad \text{for } L \text{ slowly varying}$$

A similar condition applies to the Weibull domain of attraction $\mathcal{D}(\Psi_\alpha)$ in that a distribution F belongs to the Weibull domain of attraction if and only if:

$$\bar{F}(x_F - x^{-1}) = x^{-\alpha} L(x), \qquad \text{for } L \text{ slowly varying}$$

where x_F is the finite right endpoint of the support of F.

The condition for the Gumbel domain of attraction is not as simple, a distribution F belongs to the Gumbel domain if and only if there exists a positive function $a : \mathbb{R} \to \mathbb{R}^+$ and a $t \in \mathbb{R}$ such that:

$$\lim_{x \to x_F} \frac{\bar{F}(x + ta(x))}{\bar{F}(x)} = e^{-t}$$

Throughout this report the term “heavy tailed” distribution will be used to describe any distribution in the Fréchet domain of attraction, although some of the literature refers to “heavy tailed” distributions as being the distributions that decay slower than the exponential.

At this point it will also be useful to introduce the concept of distributions that have super-heavy tails.

Definition 2.1.2 (Super Heavy Tails)**.** [FIND SUPER HEAVY TAILS DEFINITION]

Additionally, if the survival function $\bar{F}$ is slowly varying itself then F has super heavy tails i.e.

$$\lim_{x\to\infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = 1, \forall t \in \mathbb{R}^+ \implies \text{super heavy tails}$$

It is possible to gather the main three types of extremal distributions into what is called the Generalised Extreme Value (GEV) distribution [REF].

Definition 2.1.3 (Generalised Extreme Value Distribution)**.** Denoted by GEV(μ, σ, ξ) the distribution is characterised by three parameters $\mu \in \mathbb{R}$ the location, $\sigma \in \mathbb{R}^+$ the scale, and the shape $\xi \in \mathbb{R}$. It has support on $\{x \in \mathbb{R} : 1 + \xi(x-\mu)/\sigma > 0\}$ and has cdf given by:

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

The three types of extremal distribution are obtained from changing the shape parameter ξ, which corresponds to $1/\alpha$ in the definition of the domains of attraction. This change is generally to made so that increasing ξ corresponds to increasing how heavy the tails of the distribution are. So, $\xi < 0$ corresponds to the Weibull, $\xi > 0$ the Fréchet, and $\xi = 0$ the Gumbel.

While this is useful for modelling the distribution of block maxima of iid random variables, as seen in Section 1, the data in question appears to follow power law like behaviour for the bulk of the data and then changes to various different shapes above a certain threshold. For this reason, it is perhaps more appropriate to consider the distribution of threshold exceedances.

Definition 2.1.4 (Generalised Pareto Distribution)**.** The Generalised Pareto (GP) distribution can be obtained by using the GEV distribution and conditional probability such that for large enough threshold the GP distribution approximately describes the conditional distribution of threshold exceedances. More precisely, for large enough threshold u and the change of variable to $Y = X - u$:

$$\Pr(Y \leq y | Y > 0) = H(y) = \begin{cases} 1-\left(1+\frac{\xi y}{\sigma}\right)^{-1/\xi}, & y > 0, \xi \neq 0 \\ 1-\exp\left(-\frac{y}{\sigma}\right), & y > 0, \xi = 0 \end{cases}$$

Since this distribution was obtained using a GEV(μ, σ^*, ξ) the shape parameter ξ is identical in both distributions and the shape parameter σ is defined such that $\sigma = \sigma^* + \xi(u - \mu)$.

The vast majority of this theory is appropriate only for continuous data, and since the data being focused on is discrete, some results for discrete extremes should be introduced.

2.2 Discrete Extremes

3 Networks

References