INTRODUCTION TO EXTREME VALUE THEORY

Extreme Value Theory is a powerful and fairly robust framework to study the tail behaviour of a distribution asymptotically\(^1\). The main purpose of this theory is to provide asymptotic models which can model the tails of a distribution (Bystrom, 2005). EVT reduces the focus from modelling the whole distribution to modelling of the tail behaviour. Hence the symmetry assumption of the Gaussian distribution is examined directly by estimating the left and right tails separately. A critical assumption of the EVT is that extreme prices are independently and identically distributed (Nystrom and Skoglund, 2002).

Fisher and Tippett (1928) was the pioneer of EVT followed by Gnedenko (1943), Gumbel (1958), Balkema and de Haan (1972) and Pickands (1975). Most recently, Embrechts et al. (1997), Reiss and Thomas (1997), Beirlant et al. (1996) and McNeil and Frey (2001) applied EVT to insurance problems. Bystrom (2005) applied EVT to investigate the tails of the price distribution of hourly electricity spot prices in NordPool in an attempt to capture the extreme price behaviour better than conventional electricity price models. Rozario (2002) derived Value-at-risk (VaR) for Victorian half-hourly electricity returns using a threshold based EVT model whereas Chan and Gray (2006) used EVT to measure VaR for daily spot electricity prices in five different electricity markets around the globe including Victoria.

There are two commonly used ways in identifying extremes in the data; block maxima and the threshold methods.

The limit law for the block maxima with the size of the subsample is given by the following theorem of Fisher and Tippett (1928) and Gnedenko (1943);

---

\(^1\) Asymptotic theory is a generic framework for assessment of properties of estimators and statistical tests. Within this framework it is typically assumed that the sample size grows indefinitely, and the properties of statistical procedures are evaluated in the limit as sample size approaches to infinity. EVT has a fundamental role when modelling the maxima of a random variable. This role is similar to the role of Central Limit Theorem when modelling the sums of random variables. In both cases, the theory states what the limiting distributions are.
Let \((X_n)\) be a sequence of i.i.d. random variables. If there exists constants \(c_n > 0, d_n \in \mathbb{R}\) and some non-degenerate distribution function \(H\) such that \(\frac{M_n - d_n}{c_n} \to H\), then \(H\) belongs to one of following three extreme value distributions:

Frechet, \(\Phi_\alpha(x) = \begin{cases} \frac{1}{x}, & x \leq 0 \\ e^{-(-x)^{-\alpha}}, & x > 0 \end{cases} \)

Weibull, \(\Phi_\alpha(x) = \begin{cases} e^{(-x)^{\alpha}}, & x \leq 0 \\ 1, & x > 0 \end{cases} \)

Gumbel, \(\Lambda(x) = e^{-(-x)^{-\alpha}}, x \in \mathbb{R} \)

for \(\alpha > 0\)

Von Mises (1936) and Jenkinson (1955) showed one-parameter generalisations of these standard distributions as;

\[ H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-1/\xi}}, & \text{if } \xi \neq 0 \\ e^{-e^{-x}}, & \text{if } \xi = 0 \end{cases} \]

This generalisation is known as Generalised Extreme Value (GEV) distribution and obtained by setting \(\xi = \alpha^{-1}\) for the Frechet distribution, \(\xi = -\alpha^{-1}\) for the Weibull distribution and by interpreting the Gumbel distribution as the limit case for \(\xi = 0\).

Secondly, the Peaks-over-Threshold (POT) method is a threshold method where in the absence of a known distribution function \((F)\) of a random variable \((X)\), the interest lies in estimating the distribution function \(F(u)\) of values of \(x\) above a certain threshold \((u)\).

The distribution function \(F(u)\) is called the conditional excess distribution function (CEDF) and is defined as;

\[ F_u(y) = P(X - u \leq y|x > u), 0 \leq y \leq x_F - u \]

where \(X\) is a random variable, \(u\) is a given threshold, \(y = x - u\) are the excesses and \(x_F \leq \infty\) is the right end point of \(F\). This can be written in terms of \(F\) as;
The realizations of the random variable $X$ lie mainly between 0 and $u$ and therefore the estimation of $F$ in this interval are straightforward. The estimation of the portion $F(u)$ however might be difficult as there are in general very little observations in this area. Similar to the block maxima method, which provides a choice of an optimal block length, the POT method relies on a reasonable choice of threshold. A threshold value that is too low and the asymptotic theory is no longer met, too high a threshold value and one does not have enough data points to estimate the parameters in the excess distribution (Bystrom, 2005).

EVT provides a powerful result about the CEDF. This is stated by the following theorem of Pickands (1975), Balkema and de Haan (1974):

For a large class of underlying distribution function ($F$) the conditional excess distribution function $F_u(y)$, for a large $u$, is well approximated by $F_u(y) \approx G_{\xi,\sigma}(y), u \to \infty$.

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

for $y \in [0, (x_F - u)]$ if $\xi \geq 0$ and $y \in \left[0, -\frac{\sigma}{\xi}\right]$ if $\xi < 0$.

$G_{\xi,\sigma}$ is called as the Generalised Pareto Distribution (GPD).