GENERALIZED EXTREME VALUE DISTRIBUTIONS

ZHANGSHENG LAI

1 Asymptotic Models

1.1 Model Formulation

The model focuses on the statistical behaviour of $M_n = \max\{X_1, \ldots, X_n\}$ where all the X_i's are independent and identically distributed with distribution function F. The distribution of M_n can be derived exactly for all values of n

$$
\begin{aligned}
\mathbb{P}\{M_n \leq z\} &= \mathbb{P}\{X_1 \leq z, \ldots X_n \leq z\} \\
&= \prod_{i=1}^{n} \mathbb{P}\{X_i \leq z\} \\
&= \{F(z)\}^n
\end{aligned}
\tag{1.1}
$$

Although we know how to find the distribution of M_n the maximum of iid random variables, not knowing what F makes knowing the above not helpful. We could utilise standard statistical techniques like maximum likelihood to get an estimate $\widehat{F}$ from the observed data the substitute into (1.1). However small errors in the estimate of F can lead to substantial errors in F^n.

The approach we are going to look at here is to accept that F is unknown, instead of estimating F to estimate F^n, we find an estimate of F^n directly, which can be estimated using extreme data only. The idea is similar to the usual method of approximating the distribution of sample means by the normal distribution. So essentially we are doing the extreme value analogue of the central limit theory.

Observe that for a distribution function F with upper end-point z^+, i.e. $F(z^+) = 1$ for any $z < z^+$, $F^n(z) \to 0$ as $n \to \infty$, thus M_n degenerates to a point mass on z^+. To avoid this problem, we do a linear renormalization of the variable M_n:

$$
M_n^* = \frac{M_n - b_n}{a_n}
$$

for a sequence of constants $\{a_n > 0\}$ and $\{b_n\}$. By choosing appropriate $\{a_n\}$ and $\{b_n\}$ it stabilizes the location and scale of M_n^* as n grows avoiding problems of degeneracy. Thus we seek limit distributions of M_n^* instead of M_n with appropriate choices of $\{a_n\}$ and $\{b_n\}$.

1.2 Extremal Types Theorem

Theorem 1.1 (Fisher-Tippett-Gnedenko). If there exists sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$
\mathbb{P}\{M_n - b_n/a_n \leq z\} \to G(z) \quad \text{as } n \to \infty
$$

where G is a non-degenerate distribution function, then G belongs to one of the following families:

$$
\begin{aligned}
\mathbb{I}: \quad G(z) &= \exp\left\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}, \quad -\infty < z < \infty \\
\mathbb{II}: \quad G(z) &= \begin{cases} 0, & z \leq b \\ \exp\left\{-\left(\frac{z-b}{a}\right)^{-\alpha}\right\}, & z > b \end{cases} \\
\mathbb{III}: \quad G(z) &= \begin{cases} \exp\left\{-\left(-\frac{z-b}{a}\right)^{\alpha}\right\}, & z \leq b \\ 1, & z > b \end{cases}
\end{aligned}
$$

for parameters $a > 0, b$ and for families $\mathbb{II}, \mathbb{III}, \alpha > 0$.

These three classes of distribution are called **extreme value distributions** with the types $\mathbb{I}, \mathbb{II}$ and $\mathbb{III}$ widely known as the **Gumbel**, **Fréchet** and **Weibull** families respectively. Theorem 1.1 implies that when M_n can be stabilized with suitable sequences $\{a_n > 0\}$ and b_n the corresponding normalized M_n^* has a limiting distribution that must be one of the three extreme distributions. It is in this sense that the theorem provides an extreme value analog of central limit theorem.

1.3 The Generalized Extreme Value Distribution

The three types of limits have different characteristic, corresponding to the different kind of tail behaviour for the distribution function F of the X_i's. We have the Gumbel to be unbounded, Fréchet bounded below and the Weibull bounded above. The density of Gumbel decays exponentially whereas it is polynomially for the Fréchet, corresponding to relative decay rates in the tail of F. Thus the different families give different representations of extreme value behaviour. Initially, extreme value theory was applied by choosing a family that is most suitable for the data we have and once the family was chosen, we assume that subsequent inferences be made based on the assumption that the choice of family is correct and do not allow for any uncertainty although there might be much uncertainty.

A much more flexible approach is to reformulate the three models into a single distribution which has the form,

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1.2}$$

defined on the set $\{x : 1 + \xi(x - \mu)/\sigma > 0\}$ where the parameters satisfy $-\infty < \mu, \xi < \infty$ and $\sigma > 0$. The above distribution is the generalized extreme value (GEV) family of distributions, where μ is the location parameter, σ is the scale parameter and ξ is the shape parameter. The type $\mathbb{II}$ and $\mathbb{III}$ classes correspond to the case where $\xi > 0$ and $\xi < 0$ respectively and the subset of the family with $\xi = 0$ is interpreted as the limit of (1.2) as $\xi \to 0$.

Theorem 1.2. If there exists sequence of constants $\{a_n\}$ and $\{b_n\}$ such that

$$\mathbb{P}\{(M_n - b_n)/a_n \le z\} \to G(z) \text{ as } n \to \infty \tag{1.3}$$

for a non-degenerate distribution function G, then G is a member of the GEV family

$$G(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1.4}$$

defined on the set $\{x : 1 + \xi(x - \mu)/\sigma > 0\}$ where the parameters satisfy $-\infty < \mu, \xi < \infty$ and $\sigma > 0$.

By interpreting the limit in Theorem 1.2 as an approximation for large values of n suggests the use of the GEV family for modelling the distribution of maxima long sequences. With this result we can solve for the normalizing constants that are unknown is practice. Given (1.3) holds

$$\mathbb{P}\{(M_n - b_n)/a_n \le z\} \approx G(z)$$

for large n. With some manipulation

$$\mathbb{P}\{(M_n - b_n)/a_n \le z\} \approx G((z - b_n)/a_n) = G^*(z)$$

where G^* is another member of the GEV family. This says that we can approximate the distribution of M_n^* by a member of the GEV for large n, we can do the same for the distribution of M_n. Since the parameters of the distribution have to be estimated anyway, it is irrelevant in practice that the parameters of the distribution of G are different from those of G^*. **Still got stuff on return level and period here!!**

1.4 Outline Proof of the Extremal Types Theorem

Definition 1.3. A distribution G is said to be **max-stable** if, for every $n = 2, 3, \ldots$ there are constants $\alpha_n > 0$ and β_n such that

$$G^n(\alpha_n z + \beta_n) = G(z)$$

Theorem 1.4. A distribution is max-stable iff it is a generalized extreme value distribution.

Proof. We shall skip the forward direction proof as it utilizes machinery that are not trivial and show the converse. Let the Fréchet $G(x)$ be given, then

$$G^n(z) = \exp\left\{-n\left(\frac{z-b}{a}\right)^{-\alpha}\right\}$$

$$= \exp\left\{-\left(\frac{z-b}{an^{1/\alpha}}\right)^{-\alpha}\right\}$$

$$= G(\alpha_n z + \beta_n)$$

where $\alpha_n = n^{-1/\alpha}$ and $\beta_n = 1 + n^{-1/\alpha}$. In a similar fashion we can prove it for Weibull. For Gumbel, we have

$$G^n(z) = \exp\left\{-n\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}$$

$$= \exp\left\{-\exp\left(\log n\right) \cdot \exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}$$

$$= \exp\left\{-\exp\left[\log n - \left(\frac{z-b}{a}\right)\right]\right\}$$

$$= G(\alpha_n z + \beta_n)$$

where $\alpha_n = 1$ and $\beta_n = -a\log n$. $\qquad\qquad\square$

Theorem 1.4 is used directly in the proof of extremal types theorems. We start be considering M_{nk}, the maximum random variable in a sequence of $n \times k$ variables for some large value of n. This can be regarded as the maximum of a single sequence of length $n \times k$ or as the maximum of k maxima, each of which is a maximum of n observations. Let's assume that the limit distribution of $(M_n - b_n)/a_n$ is G, thus for sufficiently large n

$$\mathbb{P}\{(M_n - b_n)/a_n\} \approx G(z)$$

Hence for any integer k we also have

$$\mathbb{P}\{(M_{nk} - b_{nk})/a_{nk}\} \approx G(z)$$

but now we recall the definitions of M_n and M_{nk}, we have

$$G\left(\frac{z - b_{nk}}{a_{nk}}\right) \approx \mathbb{P}\{(M_{nk} - b_{nk})/a_{nk} \leq z\} = [\mathbb{P}\{(M_n - b_n)/a_n \leq z\}]^k \approx G^k\left(\frac{z - b_n}{a_n}\right)$$

Therefore G and G^k are identically apart except for location and scale coefficients. Thus G is max-stable and a member of GEV by Theorem 1.4

Definition 1.5 (Domain of Attraction)**.** A distribution μ is said to be in the domain of attraction of an extreme value distribution ν (either Gumbel, Fréchet or Weibull) denoted by $\mu \in \mathcal{D}(\nu)$ if there exists $\{a_n > 0\}$ and $\{b_n\}$ such that the distribution of $(M_n - b_n)/a_n$ converges weakly to ν, where $M_n := \max_{1 \leq i \leq n} X_i$ for an independent and identically distributed sequence $X_1, X_2, \ldots$ with distribution μ.

Example. If $X_1, X_2, \ldots$ is a sequence of independent and identically distributed $\mathsf{Exp}(1)$ random variables, $F(x) = 1 - e^{-x}$ for $x > 0$, choosing $a_n = 1$ and $b_n = \log n$

$$\mathbb{P}\{(M_n - b_n)/a_n\} = F^n(z + \log n)$$
$$= \left[1 - e^{-(z + \log n)}\right]^n$$
$$= \left[1 - n^{-1}e^{-z}\right]^n \to \exp(-e^{-z}) \text{ as } n \to \infty$$

here we use the fact that $\lim_{n \to \infty} \left(1 - \frac{1}{n}\right)^n = \frac{1}{e}$. Hence with the chosen a_n and b_n, the limit of M_n converges to the Gumbel distribution as $n \to \infty$. This corresponds to $\xi = 0$ in the GEV family. ▲

Example. If $X_1, X_2, \ldots$ is a sequence of independent Fréchet variables. $F(x) = \exp(-1/x)$ for $x > 0$. Letting $a_n = n$ and $b_n = 0$

$$\mathbb{P}\{(M_n - b_n)/a_n \le z\} = F^n(nz)$$
$$= \left[\exp\{-1/nz\}\right]^n$$
$$= \exp(-1/z)$$

as $n \to \infty$, for each $z > 0$. Hence the limit in this case which is an exact result for all n since the Fréchet is max-stable is also the Fréchet distribution. This corresponds to $\xi = 1$ in the GEV family. ▲

Example. If $X_1, X_2, \ldots$ are a sequence of independent uniform $\mathsf{U}(0, 1)$ variables, $F(x) = x$ for $0 \le x \le 1$. For fixed $z < 0$, suppose $n > -z$ and let $a_n = 1/n$ and $b_n = 1$. Then

$$\mathbb{P}\{(M_n - b_n)/a_n \le z\} = F^n(n^{-1}z + 1)$$
$$= \left(1 + \frac{z}{n}\right)^n$$
$$\to e^z \text{ as } n \to \infty$$

Hence the distribution is Weibull type with $\xi = -1$ in the GEV family. ▲

1.5 Inference for the GEV distribution

1.5.1 General Considerations

1.5.2 Maximum Likelihood Estimation