# Lecture 12

## 1 Max-stable Distributions

Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. real-valued random variables. It is often interesting to study

$$M_n := \max_{1\le i\le n} X_i.$$

For example, $(X_i)_{i\in\mathbb{N}}$ could be the size of earthquakes measured over time, or the severity of financial market crashes, etc. Of course in reality, the observations $(X_i)_{i\in\mathbb{N}}$ typically have correlations, nevertheless, the i.i.d. assumption provides a good starting point for the analysis. Our first observation is that, although each $X_i$ is typically of order 1, with more and more observations, we are more and more likely to observe extreme events; in other words, $M_n$ typically tends to $\infty$ in probability as $n\to\infty$. The fundamental question is

- Do there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that $\frac{M_n-b_n}{a_n}$ converges in distribution to a non-trivial limit? What is $a_n$ and $b_n$, and what are the possible limiting distributions?

Surprisingly, there are only three types of limiting distributions, the **Gumbel**, **Fréchet**, and **Weibull** distributions, known collectively as the **extreme value distributions**. This fundamental result is known as the *Extremal Types Theorem*, discovered first by Fisher and Tippett (1928), and later proved in full generality by Gnedenko (1943). In fact the Extremal Types Theorem holds for dependent $(X_i)_{i\in\mathbb{N}}$ as well under suitable conditions. Two standard references on extreme value theory are [1] and [2].

Before we state and prove the Extremal Types Theorem, we first introduce some background which brings up some parallels with the theory of stable distributions.

**Definition 1.1** **[Type of Distribution]** *Let $X$ and $Y$ be two random variables with distributions $\mu$ and $\nu$ respectively. We say that $\mu$ and $\nu$ are of the same type if there exist $a>0$ and $b\in\mathbb{R}$ such that $aX+b$ has the same distribution as $Y$.*

We recall from Lecture 11 the Convergence to Types Theorem, due to Khintchine.

**Theorem 1.2** **[Convergence to Types Theorem]** *Let $W_n$ be a sequence of random variables converging weakly to $W$, and for some $a_n>0$ and $b_n\in\mathbb{R}$, $a_nW_n+b_n \Rightarrow W'$, where both $W$ and $W'$ are non-degenerate. Then $a_n\to a$ and $b_n\to b$ for some $a>0$ and $b\in\mathbb{R}$. Equivalently, if $F_n$, $F$ and $F'$ are distribution functions with $F$ and $F'$ being non-degenerate, and there exist $a_n, a'_n>0$ and $b_n, b'_n\in\mathbb{R}$ such that $F_n(a_nx+b_n)\to F(x)$ and $F_n(a'_nx+b'_n)\to F'(x)$ at all points of continuity of $F$, respectively $F'$, then $a_n/a'_n \to a>0$, $(b_n-b'_n)/a'_n\to b\in\mathbb{R}$, and $F'(ax+b)=F(x)$ for all $x\in\mathbb{R}$.*

We first identify all non-trivial limiting distributions of $\frac{M_n-b_n}{a_n}$ with the so-called *max-stable* distributions. We then give a characterization of all max-stable distributions.

**Definition 1.3** **[Max-stable Distributions]** *A non-degenerate probability distribution $\mu$ on $\mathbb{R}$ is called max-stable, if for a sequence of i.i.d. random variables $(X_i)_{i\in\mathbb{N}}$ with distribution $\mu$*

*and for each $n \in \mathbb{N}$, there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that $\frac{M_n - b_n}{a_n}$ also has distribution $\mu$. Equivalently, $\mu$ is max-stable if its distribution function $F(x) := \mu(-\infty, x]$ satisfies: for each $n \in \mathbb{N}$, there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that*

$$F^n(a_n x + b_n) = F(x) \qquad \text{for all } x \in \mathbb{R}. \tag{1.1}$$

**Remark.** We will see in the proof of Theorem 2.1 below that the normalizing constants $a_n$ for max-stable distributions can only be of the form $a_n = n^\theta$ for some $\theta \in \mathbb{R}$, with $\theta < 0$, $\theta = 0$ and $\theta > 0$ corresponding to the three types of max-stable distribuitons: Weibull, Gumbel, and Fréchet.

**Theorem 1.4 [Max-stable Distributions are Weak Limits of Maxima]** *A non-degenerate probability distribution $\mu$ on $\mathbb{R}$ is max-stable if and only if there exist i.i.d. random variables $(X_i)_{i \in \mathbb{N}}$, $a_n > 0$ and $b_n \in \mathbb{R}$, such that the distribution of $\frac{M_n - b_n}{a_n}$ converges weakly to $\mu$.*

**Proof.** Let $Z$ be a random variable with distribution $\mu$, such that $Z$ is the weak limit of $Z_n := \frac{M_n - b_n}{a_n}$ for a sequence of i.i.d. random variables $X := (X_i)_{i \in \mathbb{N}}$ for some $a_n > 0$ and $b_n \in \mathbb{R}$. Let $(Z^{(j)})_{j \in \mathbb{N}}$ be i.i.d. copies of $Z$. For $j \in \mathbb{N}$, let $X^{(j)} := (X_i^{(j)})_{i \in \mathbb{N}}$ be i.i.d. copies of the sequence $(X_i)_{i \in \mathbb{N}}$. Let $M_n^{(j)} := \max\{X_1^{(j)}, \ldots, X_n^{(j)}\}$, and denote $Z_n^{(j)} := \frac{M_n^{(j)} - b_n}{a_n}$. Then for each $m \in \mathbb{N}$, $(Z_n^{(1)}, \cdots, Z_n^{(m)})$ converges in distribution as $n \to \infty$ to the vector of independent random variables $(Z^{(1)}, \ldots, Z^{(m)})$. Therefore by the Continuous Mapping Theorem for weak convergence,

$$\max_{1 \leq j \leq m} Z_n^{(j)} \underset{n \to \infty}{\Longrightarrow} \max_{1 \leq j \leq m} Z^{(j)}. \tag{1.2}$$

On the other hand,

$$\max_{1 \leq j \leq m} Z_n^{(j)} = \max_{1 \leq j \leq m} \frac{M_n^{(j)} - b_n}{a_n} \overset{\text{dist}}{=} \frac{M_{mn} - b_n}{a_n} = c_{m,n}\Big(\frac{M_{mn} - b_{mn}}{a_{mn}}\Big) + d_{m,n} = c_{m,n} Z_{mn} + d_{m,n},$$

where

$$c_{m,n} = \frac{a_{mn}}{a_n} \qquad \text{and} \qquad d_{m,n} = \frac{b_{mn} - b_n}{a_n}.$$

Since $Z_{mn} \Rightarrow Z$ and $c_{m,n} Z_{mn} + d_{m,n} \Rightarrow \max_{1 \leq j \leq m} Z^{(j)}$ as $n \to \infty$, where both limits are non-degenerate, we can apply the Convergence to Types Theorem, Theorem 1.2, to conclude that $c_{m,n} \to c_m > 0$ and $d_{m,n} \to d_m$ as $n \to \infty$, and

$$Z \overset{\text{dist}}{=} \frac{\max_{1 \leq j \leq m} Z^{(j)} - d_m}{c_m},$$

and hence the distribution of $Z$ is max-stable.

Conversely if $Z$ is a random variable with max-stable distribution, and $(X_i)_{i \in \mathbb{N}}$ are i.i.d. copies of $Z$, then by the definition of max-stable distributions, for each $n \in \mathbb{N}$, there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that $\frac{M_n - b_n}{a_n} \overset{\text{dist}}{=} Z$, which certainly converges in distribution to $Z$. ∎

## 2 Extremal Types Theorem

We are now ready to state and prove the Extremal Types Theorem.

**Theorem 2.1 [Extremal Types Theorem]** *Every max-stable distribution $\mu$ is of extreme value type, namely that it is of the same type as one of the following distributions:*

(1) **[Gumbel]** $\nu(-\infty, x] = e^{-e^{-x}}$ *for all* $x \in \mathbb{R}$*;*

(2) **[Fréchet]** $\nu(-\infty, x] = 0$ *if* $x \le 0$ *and* $\nu(-\infty, x] = e^{-x^{-\alpha}}$ *for some* $\alpha > 0$ *if* $x > 0$*;*

(3) **[Weibull]** $\nu(-\infty, x] = e^{-(-x)^{\alpha}}$ *for some* $\alpha > 0$ *if* $x \le 0$*, and* $\nu(-\infty, x] = 1$ *if* $x > 0$*.*

*Conversely, every distribution of extreme value type is max-stable.*

**Remark 2.2** Theorems 1.4 and 2.1 together imply that every non-degenerate limit of $\frac{M_n - b_n}{a_n}$, where $M_n := \max_{1 \le i \le n} X_i$ for a sequence of i.i.d. random variables $(X_i)_{i \in \mathbb{N}}$, is of extreme value type. Conversely, every distribution $\mu$ of extreme value type is the weak limit of $\frac{M_n - b_n}{a_n}$ for some $a_n > 0$ and $b_n \in \mathbb{R}$, where we can choose $(X_i)_{i \in \mathbb{N}}$ to have distribution $\mu$.

**Remark 2.3** Note that if $\nu$ has the Gumbel distribution, then $\nu(x, \infty) \approx e^{-x}$ as $x \to \infty$, while $\nu(-\infty, -x) \approx e^{-e^{x}}$ decays double exponentially fast as $x \to \infty$; If $\nu$ has the Fréchet distribution, then $\nu(-\infty, 0] = 0$ and $\nu(x, \infty) \approx x^{-\alpha}$ as $x \to \infty$; If $\nu$ has the Weibull distribution, then $\nu[0, \infty) = 0$, and $\nu(-x, 0) \approx x^{\alpha}$ as $x \downarrow 0$.

**Proof of Theorem 2.1.** If $F$ is the distribution function of either a Gumbel, Fréchet, or Weibull type distribution, then it is easy to check that $F$ satisfies (1.1) for some $a_n > 0$ and $b_n \in \mathbb{R}$, and hence distributions of extreme value type are max-stable.

Conversely, let $F$ be the distribution function of a max-stable distribution $\mu$. Then by (1.1), for all $s > 0$ and $n \in \mathbb{N}$,

$$F^{[ns]}(a_{[ns]}x + b_{[ns]}) = F(x) \qquad \text{for all } x \in \mathbb{R},$$

while

$$F^{[ns]}(a_n x + b_n) = \big(F^n(a_n x + b_n)\big)^{[ns]/n} = F^{[ns]/n}(x) \underset{n\to\infty}{\longrightarrow} F^s(x) \qquad \text{for all } x \in \mathbb{R}.$$

Since $F^{[ns]}$, $F$ and $F^s$ are all non-degenerate distribution functions, we can apply the Convergence to Types Theorem, Theorem 1.2, to conclude that

$$\frac{a_{[ns]}}{a_n} \underset{n\to\infty}{\longrightarrow} a_s > 0, \qquad \frac{b_{[ns]} - b_n}{a_n} \underset{n\to\infty}{\longrightarrow} b_s \in \mathbb{R}, \qquad \text{and} \quad F^s(a_s x + b_s) = F(x). \tag{2.3}$$

Therefore for any $s, t > 0$,

$$\begin{aligned} F(x) &= F^{st}(a_{st}x + b_{st}) \\ &= \big(F^s(a_{st}x + b_{st})\big)^t = \Big(F\Big(\frac{a_{st}}{a_s}x + \frac{b_{st} - b_s}{a_s}\Big)\Big)^t = F\Big(\frac{a_{st}}{a_s a_t}x + \frac{b_{st} - b_s - a_s b_t}{a_s a_t}\Big). \end{aligned}$$

It is an easy exercise to show that

**Exercise 2.4** If $F$ is the distribution function of a non-degenerate distribution, and for some $a > 0$ and $b \in \mathbb{R}$, $F(ax + b) = F(x)$ for all $x \in \mathbb{R}$, then $a = 1$ and $b = 0$.

Therefore we have

$$a_{st} = a_s a_t \qquad \text{and} \qquad b_{st} = b_s + a_s b_t = b_t + a_t b_s \qquad \text{for all } s, t > 0. \tag{2.4}$$

In particular, $f(x) := \log a_{e^x}$ satisfies the *Cauchy Functional Equation*

$$f(x + y) = f(x) + f(y). \tag{2.5}$$

We observe that $a_s$ and $b_s$ are continuous in $s > 0$, since (2.3) implies that for any $s_0 > 0$,

$$F^s(a_s x + b_s) \underset{s\to s_0}{\longrightarrow} F(x) \qquad \text{and} \qquad F^s(a_{s_0} x + b_{s_0}) \underset{s\to s_0}{\longrightarrow} F(x),$$

and hence $a_s \to a_{s_0}$ and $b_s \to b_{s_0}$ as $s \to s_0$ by the Convergence to Types Theorem and Exercise 2.4. Therefore $f(x) = \log a_{e^x}$ is a continuous solution of (2.5). It is left as an exercise to show that any solution of (2.5) which is continuous (and hence bounded on finite intervals) must be of the form

$$f(x) = \theta x \qquad \text{for some } \theta \in \mathbb{R}.$$

Therefore

$$a_t = t^\theta. \tag{2.6}$$

We consider the three cases (a): $\theta = 0$, (b): $\theta > 0$, (c): $\theta < 0$.

If $\theta = 0$, then $a_\cdot \equiv 1$, and (2.4) gives

$$b_{st} = b_s + b_t.$$

The function $g(x) := b_{e^x}$ is again a continuous solution of the Cauchy functional equation (2.5), and hence $g(x) = cx$ and $b_t = c \log t$ for some $c \in \mathbb{R}$. By (2.3), we have

$$F^s(x + c\log s) = F(x) \qquad \text{for all } s > 0,\ x \in \mathbb{R}. \tag{2.7}$$

If $c = 0$, then $F(x) \in \{0, 1\}$ for all $x \in \mathbb{R}$, which implies that $F$ is degenerate and contradicts our assumption. Therefore $c \neq 0$. Since $F$ is non-degenerate, $F(x_0) \in (0, 1)$ for some $x_0 \in \mathbb{R}$, and we may assume without loss of generality that $x_0 = 0$. Then (2.7) gives

$$F(y) = F(0)^{e^{-y/c}} = e^{-c' e^{-y/c}} \qquad \text{with} \quad c' = -\log F(0) \in (0, \infty),$$

where $c > 0$ since $F$ is non-decreasing. A linear change of variable then gives the Gumbel distribution function $e^{-e^{-x}}$.

Now we assume $a_t = t^\theta$ with $\theta \neq 0$. By (2.4),

$$\frac{b_s}{1 - a_s} = \frac{b_t}{1 - a_t}.$$

Fix a $t_0$ such that $a_{t_0} \in (0, 1)$. Then

$$b_t = \frac{1 - t^\theta}{1 - t_0^\theta}\, b_{t_0} =: c(1 - t^\theta).$$

By (2.3), we have

$$F(x) = F^s(s^\theta x + c(1 - s^\theta)) = F^s(s^\theta (x - c) + c).$$

Let $y = x - c$ and $G(y) := F(x + c)$, which is a distribution function of the same type as $F$, we then obtain

$$G(y) = G^s(s^\theta y) \qquad \text{for all} \quad s > 0,\ y \in \mathbb{R}. \tag{2.8}$$

If $\theta > 0$, then by letting $s \uparrow \infty$, we see that $G(y) = 0$ for all $y < 0$. Since $G$ is non-degenerate, $G(y_0) \in (0, 1)$ for some $y_0 > 0$, which we may assume without loss of generality to be $y_0 = 1$. From (2.8), we then obtain

$$G(y) = G(1)^{y^{-1/\theta}} = e^{-c' y^{-1/\theta}} \qquad \text{for all} \quad y \geq 0,$$

where $c' = -\log G(1) \in (0, \infty)$. A linear change of variable then gives the Fréchet distribution function $1_{\{x>0\}} e^{-x^{-\alpha}}$ for some $\alpha = 1/\theta > 0$.

The case for $\theta < 0$ is similar, which leads to the Weibull distribution. ∎

# 3 Domain of Attraction

**Definition 3.1 [Domain of Attraction]** *A distribution $\mu$ is said to be in the domain of attraction of an extreme value type distribution $\nu$ (either Gumbel, Fréchet, or Weibull), denoted by $\mu \in D(\nu)$, if there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that the distribution of $\frac{M_n - b_n}{a_n}$ converges weakly to $\nu$, where $M_n := \max_{1 \le i \le n} X_i$ for an i.i.d. sequence $(X_i)_{i \in \mathbb{N}}$ with distribution $\mu$.*

The complete characterization of the domain of attraction of the Gumbel distribution is somewhat complicated and involves refined notions of regularly varying functions. See e.g. [1, 2]. We leave it as an advanced exercise to verify the following sufficient conditions for a distribution to be in the domain of attraction of the Gumbel distribution.

**Exercise 3.2 [Effective Criteria for Attraction to Gumbel]** Let $F(x) := 1 - e^{-h(x)}$ be the distribution function of a probability measure $\mu$. Assume that $h$ is differentiable. Then $\mu$ is in the domain of attraction of the Gumbel distribution if either of the following conditions are satisfied: (i) $h'(x) = x^{\alpha-1} L(x)$ for some $\alpha > 0$ and a slowly varying (at $\infty$) function $L$; (ii) $h(x) = x^{\alpha} L(x)$ for some $\alpha > 0$ and a slowly varying function $L$, and $h'(x)$ is monotone on $(x_0, \infty)$ for some $x_0 > 0$.

In particular, the exponential and the Gaussian distributions are both in the domain of attraction of the Gumbel distribution. Let us work out the norming constants for the exponential distribution and see how the Gumbel distribution arises.

**Example 3.3 [Exponential Distribution]** Let $F(x) = 1 - e^{-x}$, which is the distribution function of the exponential distribution with mean 1. Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. mean 1 exponential random variables. Finding $a_n > 0$ and $b_n \in \mathbb{R}$ such that $\frac{M_n - b_n}{a_n}$ converges in distribution to a non-trivial limit is equivalent to having

$$F^n(a_n x + b_n) \underset{n \to \infty}{\longrightarrow} G(x) \qquad \text{for all} \;\; x \in \mathbb{R} \tag{3.9}$$

for some non-degenerate distribution function $G$, for which we may assume without loss of generality that $G(0) \in (0, 1)$. In particular, for $x \in \mathbb{R}$ with $G(x) \neq 0$, this gives

$$\log F^n(a_n x + b_n) = n \log\left(1 - e^{-a_n x - b_n}\right) \underset{n \to \infty}{\longrightarrow} \log G(x),$$

and hence by Taylor expansion,

$$n e^{-a_n x - b_n} \underset{n \to \infty}{\longrightarrow} -\log G(x). \tag{3.10}$$

For the above convergence to hold for $x = 0$, we may take

$$b_n := \log n - \log(-\log G(0))$$

so that $n e^{-b_n} = -\log G(0)$. Setting $x = x_0$ in (3.10) for an $x_0 > 0$ with $G(x_0) \in (0, 1)$ gives

$$e^{-a_n x_0} \underset{n \to \infty}{\longrightarrow} \frac{\log G(x_0)}{\log G(0)}. \tag{3.11}$$

Therefore we can choose

$$a_n := -\frac{1}{x_0} \log \frac{\log G(x_0)}{\log G(0)} =: C > 0.$$

Substituting this choice of $b_n$ and $a_n$ back into (3.10) then gives

$$G(x) = e^{\log G(0) e^{-Cx}},$$

which is of Gumbel type. Since $\log G(0)$ and $\log G(x_0)$ can be arbitrary constants in $(-\infty, 0)$ by changing the centering and scaling of the limit distribution, in retrospect, we can choose $b_n = \log n$ and $a_n = 1$, which give $G(x) = e^{-e^{-x}}$ in (3.9).

We now give necessary and sufficient conditions for a distribution to be in the domain of attraction of the Fréchet, respectively Weibull distribution.

**Theorem 3.4** **[Domain of Attraction of Fréchet Distribution]** *Let $\alpha > 0$, and let $\Phi_\alpha$ be the distribution function of a Fréchet distribution with index $\alpha$, i.e., $\Phi_\alpha(x) := 0$ if $x \leq 0$ and $\Phi_\alpha(x) = e^{-x^{-\alpha}}$ if $x > 0$. Then a distribution $\mu$ with distribution function $F$ is in the domain of attraction of $\Phi_\alpha$ if and only if $\mu(x, \infty) = x^{-\alpha} L(x)$ for some slowly varying function $L$. In this case,*

$$F^n(a_n x) \to \Phi_\alpha(x) \qquad \textit{for all} \;\; x \in \mathbb{R},$$

*where*

$$a_n := \Big(\frac{1}{1-F}\Big)^{\leftarrow}(n),$$

*with $f^{\leftarrow}(y) := \inf\{s : f(x) \geq y\}$ being the right inverse of a non-decreasing function $f$.*

**Remark.** It is easy to see that, $a_n$ is chosen such that given $n$ i.i.d. random variables $(X_i)_{1 \leq i \leq n}$ with distribution $\mu$, the number of $X_i$'s that exceed $a_n$ is a mean 1 Poisson random variable. In fact, if we denote by $X'_1 > X'_2 > \cdots$ an ordering of $X_1, \ldots, X_n$, then as a random measure on $\mathbb{R}$, $\delta_{X'_1/a_n} + \delta_{X'_2/a_n} + \cdots$ converges in distribution to a Poisson Point Process on $(0, \infty)$ with intensity measure $\frac{\alpha}{x^{\alpha+1}} \mathrm{d}x$.

**Theorem 3.5** **[Domain of Attraction of Weibull Distribution]** *Let $\alpha > 0$, and let $\Psi_\alpha$ be the distribution function of a Weibull distribution with index $\alpha$, i.e., $\Phi_\alpha(x) = e^{-(-x)^\alpha}$ if $x < 0$ $\Psi_\alpha(x) := 1$ if $x \geq 0$. Then a distribution $\mu$ with distribution function $F$ is in the domain of attraction of $\Psi_\alpha$ if and only if there exists $x_0 \in \mathbb{R}$, such that $\mu(-\infty, x_0) = 1$ and $\mu(x_0 - \epsilon, x_0) = \epsilon^\alpha L(1/\epsilon)$ for some slowly varying function $L$ as $\epsilon \downarrow 0$. In this case, we can set*

$$\gamma_n := \Big(\frac{1}{1-F}\Big)^{\leftarrow}(n),$$

*and then*

$$F^n(x_0 + (x_0 - \gamma_n)x) \underset{n\to\infty}{\longrightarrow} \Psi_\alpha(x) \qquad \textit{for all} \;\; x \in \mathbb{R}.$$

**Remark.** Similar to the remark following Theorem 3.4, one can show that given i.i.d. random variables $(X_i)_{1 \leq i \leq n}$ with distribution $\mu$, and $X'_1 > X'_2 > \cdots$ denotes their reordering, then $\sum_{i=1}^n \delta_{(X'_i - x_0)/(x_0 - \gamma_n)}$ converges in distribution to a Poisson Point Process on $(-\infty, 0)$ with intensity measure $\alpha |x|^{\alpha-1} \mathrm{d}x$.

For a proof of Theorems 3.4 and 3.5 and the remarks following them, see e.g. [1, 2], which rely on the theory of regularly varying functions.

## References


[1] M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and related properties of random sequences and processes.* Springer-Verlag, 1983.

[2] S. Resnick. *Extreme values, regular variation, and point processes.* Springer-Verlag, 1987.