

Fisher - Tippet - Gnedenko thm.

Thm (Fisher - Tippet - Gnedenko: classification of limit laws)

Suppose there exists sequences $a_n > 0$ and $b_n \in \mathbb{R}$ s.t.

$$F^n(a_n z + b_n) \xrightarrow{n \rightarrow \infty} G(z)$$

where $G(z)$ is a non-degenerate distribution function, then there exists $\mu \in \mathbb{R}$, $\sigma > 0$, $\xi \in \mathbb{R}$ s.t.

$$G(z) = \exp\left(-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right), \quad 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0$$

where for $\xi = 0$, the interpretation is

$$G(z) = \exp\left(-\exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right), \quad x \in \mathbb{R}.$$

Conversely, for each such G there is a non-empty class of distributions F (its "maximum domain of attraction") and normalising constants $a_n > 0$, $b_n \in \mathbb{R}$ such that $F^n(a_n z + b_n) \rightarrow G(z)$ for all continuity points z of G .

Proof requires two directions:

- Classification-direction: any non-degenerate limit law of normalised maxima must be GEV.
- Domain-of-attraction: what type of tail implies which type of GEV?

Fréchet type: (bounded from below)

For $\delta > 0$, limiting distribution converges if and only if $x_{\max} = +\infty$ and

$$\lim_{t \rightarrow \infty} \frac{1 - F(ut)}{1 - F(t)} = u^{-1/\delta} \quad \forall u > 0$$

$\Rightarrow b_n = 0$ and $a_n = F^{-1}\left(1 - \frac{1}{n}\right)$ possible sequences
 $\rightarrow$ a "heavy tailed distribution"

Gumbel ($\gamma=0$, unbounded) Convergence condition:

$$\lim_{t \rightarrow x_{\max}} \frac{1 - F(t + u\tilde{g}(t))}{1 - F(t)} = e^{-u} \quad \forall u > 0, \quad \tilde{g}(t) = \frac{\int_t^{x_{\max}} (1 - F(t))}{1 - F(t)}$$

Possible sequences: $b_n = F^{-1}(1 - \frac{1}{n})$ and $a_n = \tilde{g}(F^{-1}(1 - \frac{1}{n}))$.

Weibull ($\gamma < 0$, bounded from above with $x_{\max} < +\infty$)

Convergence condition:

$$\lim_{t \rightarrow 0^+} \frac{1 - F(x_{\max} - ut)}{1 - F(x_{\max} - t)} = u^{-1/\gamma} \quad \forall u > 0$$

(Note that $-1/\gamma > 0$ since $\gamma < 0$)

Possible sequences: $b_n = x_{\max}$ and $a_n = x_{\max} - F^{-1}(1 - \frac{1}{n})$.

GEV: Convergence of exponential dist. to Gumbel

Exp. dist. with parameter $\lambda > 0$ is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Let $X_1, \dots, X_n \sim F$ and $M_n = \max\{X_1, \dots, X_n\}$.

Then, for $x \geq 0$, $P(M_n \leq x) = (F(x))^n = (1 - e^{-\lambda x})^n$.

Now, consider

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) = P(M_n \leq a_n z + b_n) = (F(a_n z + b_n))^n.$$

We want to show that this converges to $\exp(-e^{-z})$.

For $x \geq 0$, $1 - F(x) = e^{-\lambda x}$.

$\Rightarrow$ We can write $F(a_n z + b_n) = 1 - \exp(-\lambda(a_n z + b_n))$

$$\Rightarrow (F(a_n z + b_n))^n = (1 - \exp(-\lambda(a_n z + b_n)))^n$$

If u_n small, we have $(1 - u_n)^n \approx \exp(-n u_n)$.

Set $u_n = \exp(-\lambda(a_n z + b_n))$. Then,

$$(F(a_n z + b_n))^n \approx \exp(-n \exp(-\lambda(a_n z + b_n))).$$

$$\Rightarrow \text{We need } \underbrace{n \exp(-\lambda(a_n z + b_n))}_{= n \exp(-\lambda b_n) \exp(-\lambda a_n z)} \rightarrow \exp(-z) \text{ as } n \rightarrow \infty$$

Now,

$$\exp(-\lambda a_n z) \rightarrow \exp(-z) \text{ if } \lambda a_n \rightarrow 1, \text{ or } a_n = \frac{1}{\lambda}$$

and

$$n \exp(-\lambda b_n) \rightarrow 1 \text{ if } \exp(-\lambda b_n) \rightarrow \frac{1}{n}, \text{ if } b_n = \frac{1}{\lambda} \log n. \quad \blacksquare$$

Extremal index

Let $Y_1, \dots, Y_{r_n}$ be a block of length r_n , u_n a high threshold. Let

$$N_{r_n}(u_n) = \sum_{j=1}^{r_n} \mathbb{1}\{Y_j > u_n\}$$

be the number of exceedances in the block and $Z_{r_n} = \max\{Y_1, \dots, Y_{r_n}\}$.

For large u_n , we have either

- no exceedances (no cluster), or
- one cluster of exceedances.

Then,

$$\begin{aligned} \text{mean cluster size} &\approx \mathbb{E}[N_{r_n}(u_n) \mid N_{r_n}(u_n) > 0] \\ &= \sum_{j=1}^{r_n} \mathbb{E}[\mathbb{1}\{Y_j > u_n\} \mid N_{r_n}(u_n) > 0] \end{aligned}$$

Now, by stationarity,

$$\mathbb{E}[N_{r_n}(u_n)] = \sum_{j=1}^{r_n} \mathbb{P}(Y_j > u_n) = r_n \mathbb{P}(Y_1 > u_n) = r_n (1 - F(u_n))$$

Also,

$$\mathbb{P}(N_k(u_n) > 0) = \mathbb{P}(M_k > u_n)$$

and

$$\mathbb{E}[N_{r_n}(u_n) \mid N_{r_n}(u_n) > 0] = \mathbb{E}[N_{r_n}(u_n) \mid M_k > u_n]$$

Thus, by rules of conditional expectation

$$\begin{aligned} \mathbb{E}[N_{r_n}(u_n) \mid N_{r_n}(u_n) > 0] &= \mathbb{E}[N_{r_n}(u_n) \mid M_k > u_n] \\ &= \frac{\mathbb{E}[N_{r_n}(u_n)]}{\mathbb{P}(M_k > u_n)} = \frac{r_n (1 - F(u_n))}{\mathbb{P}(M_k > u_n)} \end{aligned}$$