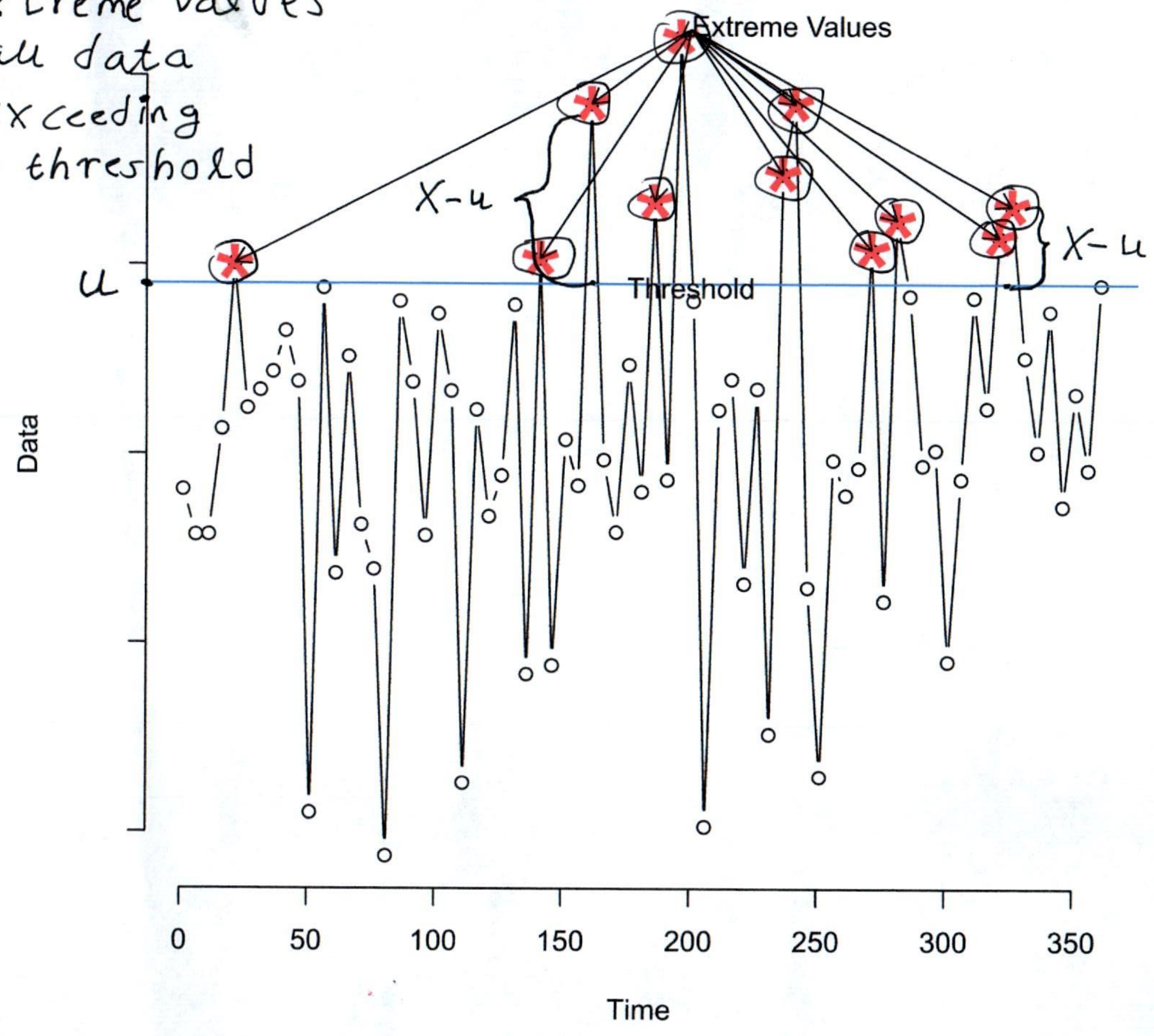
Definition 2

Extreme values

= all data
exceeding
a threshold



Definition 2 of extreme values

Suppose X = variable of interest

u = threshold.

If $X > u$ then it is an extreme value.

What is the distribution of X when $X > u$?

There is a result due to Pickands (1975) which says

$$P(\underbrace{\boxed{X-u}}_{\text{excess amount}} > x \mid \boxed{X > u})$$

X is an extreme value

$$\longrightarrow \left(1 + \xi \frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

as $u \longrightarrow w(F)$, where $F(\cdot)$ denotes the CDF of X.

Suppose u is large enough such that

$$P(X-u > x \mid X > u)$$

$$\approx \left(1 + \xi \frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad \frac{P(X-u>x,\ X>u)}{P(X>u)} \approx \left(1+\xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad \frac{P(X>u+x,\ X>u)}{P(X>u)} \approx \left(1+\xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad \frac{P(X>u+x)}{P(X>u)} \approx \left(1+\xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad P(X>u+x) \approx P(X>u)\left(1+\xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad 1-P(X\leq u+x) \approx P(X>u)\left(1+\xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad 1-F(u+x) \approx P(X>u)\left(1+\xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}$$

$\boxed{\text{Set } y = u+x}$

$$\Rightarrow \quad 1-F(y) \approx P(X>u)\left(1+\xi\frac{y-u}{\sigma}\right)^{-\frac{1}{\xi}}$$

$$\Rightarrow \quad \boxed{F(y) \approx 1-P(X>u)\left(1+\xi\frac{y-u}{\sigma}\right)^{-\frac{1}{\xi}}}$$

Generalised Pareto (GP) distribution

The GP distribution

A random variable X has the GP distribution if its CDF is

$$F(y) = 1 - P(X > u)\left(1 + \xi \frac{y-u}{\sigma}\right)^{-\frac{1}{\xi}}$$

where $1 + \xi \frac{y-u}{\sigma} > 0$, $y > u$,

$\sigma > 0$ "scale parameter"

$-\infty < \xi < \infty$ "shape parameter"

<u>Notation</u> : $X \sim GP(\sigma, \xi)$

Domain of $X \sim GP(\sigma, \xi)$

$\boxed{\xi > 0}$

$$1 + \xi \frac{y-u}{\sigma} > 0 \quad \& \quad y > u$$

$$\Longleftrightarrow \quad \xi \frac{y-u}{\sigma} > -1 \quad \& \quad y > u$$

$$\Longleftrightarrow \quad \frac{y-u}{\sigma} > -\frac{1}{\xi} \quad \& \quad y > u$$

$$\Longleftrightarrow \quad y > u - \frac{\sigma}{\xi} \quad \& \quad y > u$$

$$\Longleftrightarrow \quad y > u$$

$\Longrightarrow$ the domain $(u, +\infty)$.

$\boxed{\xi < 0}$

$$1 + \xi \frac{y-u}{\sigma} > 0 \quad \& \quad y > u$$

$$\Longleftrightarrow \quad \xi \frac{y-u}{\sigma} > -1 \quad \& \quad y > u$$

$$\Longleftrightarrow \quad \frac{y-u}{\sigma} < -\frac{1}{\xi} \quad \& \quad y > u$$

$$\Longleftrightarrow \quad y < u - \frac{\sigma}{\xi} \quad \& \quad y > u$$

$$\Longleftrightarrow \quad u < y < u - \frac{\sigma}{\xi}$$

$\Longrightarrow$ the domain $\left(u, u - \frac{\sigma}{\xi}\right)$.

$\boxed{\xi = 0}$

$1 + \xi \frac{y-u}{\sigma} > 0$ & $y > u$

$\Leftrightarrow$ $1 + 0 > 0$ & $y > u$

$\Leftrightarrow$ $y > u$

$\Rightarrow$ the domain is $(u, +\infty)$

Hence,

$$\text{Domain of } GP(\sigma, \xi) = \begin{cases} (u, \infty) & \text{if } \xi \geq 0 \\ \left(u, u - \frac{\sigma}{\xi}\right) & \text{if } \xi < 0 \end{cases}$$

The $GP(\sigma, 0)$ distribution

$$F(y) = \lim_{\xi \to 0} \left[1 - P(X > u) \left(1 + \xi \frac{y-u}{\sigma} \right)^{-\frac{1}{\xi}} \right]$$

$$= 1 - P(X > u) \lim_{\xi \to 0} \left(1 + \xi \frac{y-u}{\sigma} \right)^{-\frac{1}{\xi}}$$

$$= 1 - P(X > u) \lim_{\xi \to 0} \left(1 + \frac{\frac{y-u}{\sigma}}{\frac{1}{\xi}} \right)^{-\frac{1}{\xi}}$$

$\boxed{\text{Set } m = \frac{1}{\xi}}$

$$= 1 - P(X > u) \lim_{m \to \infty} \left(1 + \frac{\frac{y-u}{\sigma}}{m} \right)^{-m}$$

$$= 1 - P(X > u) \left[\lim_{m \to \infty} \left(1 + \frac{\frac{y-u}{\sigma}}{m} \right)^{m} \right]^{-1}$$

$$\boxed{\lim_{m \to \infty} \left(1 + \frac{z}{m} \right)^{m} = e^{z}}$$

$$= 1 - P(X > u) \left[e^{\frac{y-u}{\sigma}} \right]^{-1}$$

$$= 1 - P(X > u) \, e^{-\frac{y-u}{\sigma}}$$

= CDF of a shifted exponential distribution

PDF of GP(σ, ξ) distribution

$$f(y) = \frac{d}{dy} F(y)$$

$$= \frac{d}{dy} \left[1 - P(X > u) \left(1 + \xi \frac{y-u}{\sigma}\right)^{-\frac{1}{\xi}} \right]$$

$$= P(X > u) \cdot \frac{1}{\sigma} \cdot \left(1 + \xi \frac{y-u}{\sigma}\right)^{-\frac{1}{\xi}}$$

Quantile of $GP(\sigma, \xi)$ distribution

If X is a RV its p^{th} quantile is defined by

$$P(X \leq x_p) = p.$$

If $X \sim GP(\sigma, \xi)$ then

$$1 - P(X > u)\left(1 + \xi \frac{x_p - u}{\sigma}\right)^{-\frac{1}{\xi}} = p$$

$$\Leftrightarrow \quad \left(1 + \xi \frac{x_p - u}{\sigma}\right)^{-\frac{1}{\xi}} = \frac{1-p}{P(X > u)}$$

$$\Leftrightarrow \quad 1 + \xi \frac{x_p - u}{\sigma} = \left[\frac{1-p}{P(X > u)}\right]^{-\xi}$$

$$\Leftrightarrow \quad \boxed{x_p = u + \frac{\sigma}{\xi}\left\{\left[\frac{1-p}{P(X > u)}\right]^{-\xi} - 1\right\}}$$

In particular,

$$x_{\frac{1}{2}} = \text{Median}(X)$$

$$= u + \frac{\sigma}{\xi}\left\{\left[\frac{1}{2P(X > u)}\right]^{-\xi} - 1\right\}.$$

Return level of $GP(\sigma, \xi)$ distribution

x_T is the level that will be exceeded on average once in every T years. By definition,

$$P(X > x_T) = \frac{1}{mT} \quad \ldots \quad (++)$$

where m denotes the average number of extremes per year.

$$(++) \iff P(X \leq x_T) = 1 - \frac{1}{mT}$$

$$\iff 1 - P(X > u)\left(1 + \xi \frac{x_T - u}{\sigma}\right)^{-\frac{1}{\xi}} = 1 - \frac{1}{mT}$$

$$\iff \boxed{x_T = u + \frac{\sigma}{\xi}\left\{\left[\frac{1}{mTP(X>u)}\right]^{-\xi} - 1\right\}}$$

Estimation of $GP(\sigma, \xi)$ distribution

Suppose $x_1, \ldots, x_n$ are IID data from $GP(\sigma, \xi)$. The likelihood function is

$$L(\sigma, \xi) = \prod_{i=1}^{n} \left[\frac{P(X > u)}{\sigma} \left(1 + \xi \frac{x_i - u}{\sigma}\right)^{-\frac{1}{\xi} - 1} \right]$$

$$= \frac{[P(X > u)]^n}{\sigma^n} \left[\prod_{i=1}^{n} \left(1 + \xi \frac{x_i - u}{\sigma}\right) \right]^{-\frac{1}{\xi} - 1}.$$

The log likelihood function is

$$\log L(\sigma, \xi) = n \log P(X > u) - n \log \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{n} \log \left(1 + \xi \frac{x_i - u}{\sigma}\right).$$

The MLEs of σ and ξ are the simultaneous solutions of

$$\frac{\partial \log L}{\partial \sigma} = 0$$

and

$$\frac{\partial \log L}{\partial \xi} = 0.$$

MLE equations for the GP distribution

The MLEs of σ and ξ are the simultaneous solutions of

$$\begin{aligned}\frac{\partial \log L}{\partial \sigma} &= -\frac{n}{\sigma} + \frac{1+\xi}{\sigma^2}\sum_{i=1}^{n}(x_i - u)\left(1+\xi\frac{x_i - u}{\sigma}\right)^{-1} \\ &= 0, \qquad (1)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \log L}{\partial \xi} &= \frac{1}{\xi^2}\sum_{i=1}^{n}\log\left(1+\xi\frac{x_i - u}{\sigma}\right) \\ &\quad -\frac{1+\xi}{\xi\sigma}\sum_{i=1}^{n}(x_i - u)\left(1+\xi\frac{x_i - u}{\sigma}\right)^{-1} \qquad (2) \\ &= 0.\end{aligned}$$

The MLEs of σ and ξ are the simultaneous solutions of (1) and (2).

<u>fpot(•)</u> in R can compute the MLEs of σ and ξ.