Lecture 11

Extreme Value Analysis

Readings: An Introduction to Statistical Modeling of Extreme Values, Stuart Coles, 2001

MATH 8090 Time Series Analysis
October 26 & October 28, 2021

Whitney Huang
Clemson University

Agenda

1. **Motivation**
2. **Extreme Value Theorem & Block Maxima Method**
3. **Peaks–Over–Threshold (POT) Method**
4. **Non-stationary Extreme Value Analysis**

Outline

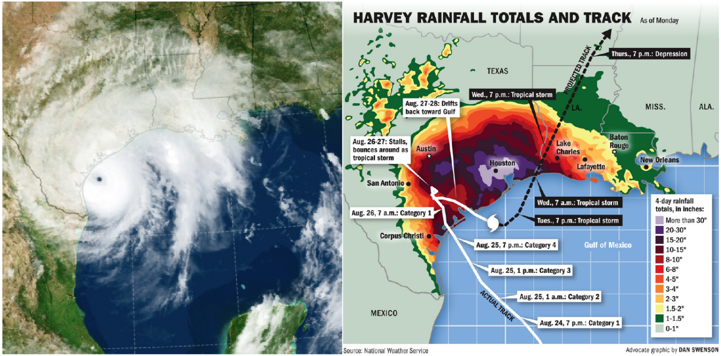
1 **Motivation**

2 Extreme Value Theorem & Block Maxima Method

3 Peaks–Over–Threshold (POT) Method

4 Non-stationary Extreme Value Analysis

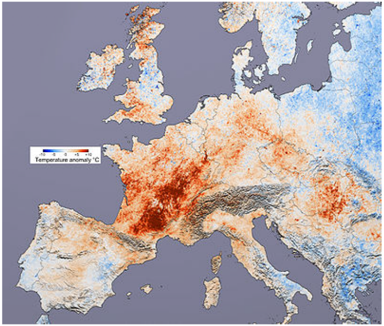
Extreme Rainfall During Hurricane Harvey



Source: NASA (Left); National Weather Service (Right)

- *"A storm forces Houston, the limitless city, to consider its limits"* – The New York Times (8.31.17)

Environmental Extremes: Heatwaves, Storm Surges, etc.



- **Heat wave**: The 2003 European heat wave led to the hottest summer on record in Europe since 1540 that resulted in at least 30,000 deaths

- **Storm Surge**: Hurricane Katrina produced the highest storm surge ever recorded (27.8 feet) on the U.S. coast

Why Study Extremes?

Although infrequent, extremes usually have large impact.

Goal: to quantify the tail behavior $\Rightarrow$ often requires extrapolation.

Applications:

- Hydrology: flooding
- Climate: temperature, precipitation, wind, $\cdots$
- Finance
- Insurance/reinsurance
- Engineering: structural design, reliability

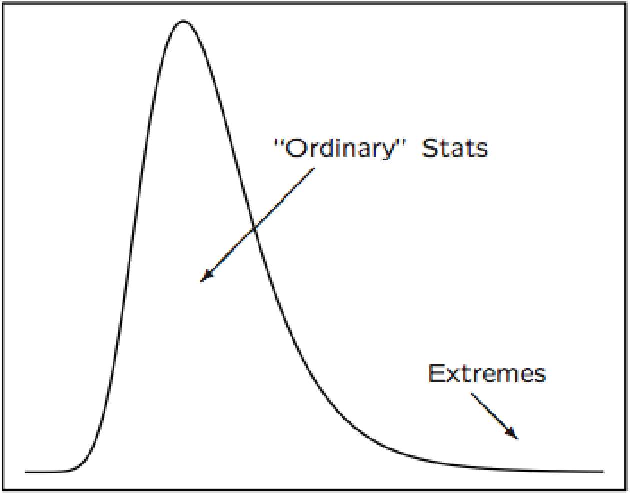
Some Scientific Questions

- How to estimate the magnitude of extreme events (e.g. 100-year rainfall)?
- How extremes vary in space?
- How extremes may change in future climate conditions?

Outline

1. Motivation
2. **Extreme Value Theorem & Block Maxima Method**
3. Peaks–Over–Threshold (POT) Method
4. Non-stationary Extreme Value Analysis

Usual vs Extremes



Probability Framework

Let $X_1, \cdots, X_n \overset{i.i.d.}{\sim} F$ and define $M_n = \max\{X_1, \cdots, X_n\}$
Then the distribution function of M_n is

$$\begin{aligned} \mathbb{P}(M_n \le x) &= \mathbb{P}(X_1 \le x, \cdots, X_n \le x) \\ &= \mathbb{P}(X_1 \le x) \times \cdots \times \mathbb{P}(X_n \le x) = F^n(x) \end{aligned}$$

Remark

$$F^n(x) \overset{n \to \infty}{=\!=} \begin{cases} 0 & \text{if } F(x) < 1 \\ 1 & \text{if } F(x) = 1 \end{cases}$$

$\Rightarrow$ the limiting distribution is degenerate.

Asymptotic: Classical Limit Laws

Recall the Central Limit Theorem:

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1),$$

where $S_n = \sum_{i=1}^{n} X_i$
$\Rightarrow$ rescaling is the key to obtain a non-degenerate distribution

Asymptotic: Classical Limit Laws

Recall the Central Limit Theorem:

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0,1),$$

where $S_n = \sum_{i=1}^{n} X_i$
$\Rightarrow$ rescaling is the key to obtain a non-degenerate distribution

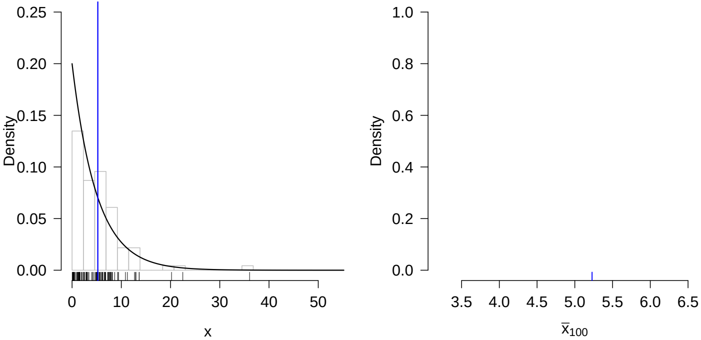
Question: Can we get the limiting distribution of

$$\frac{M_n - b_n}{a_n}$$

for suitable sequence $\{a_n\} > 0$ and $\{b_n\}$?

CLT in Action

1. Generate 100 (n) random numbers from an Exponential distribution (population distribution)
2. Compute the sample mean of these 100 random numbers
3. Repeat this process 120 times



Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \xrightarrow{d} \mathsf{G}(x)$$

Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \xrightarrow{d} \mathsf{G}(x)$$

then G must be the same type of the following form:

$$\mathsf{G}(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi(\frac{x - \mu}{\sigma})\right]_+^{\frac{-1}{\xi}}\right\}$$

where $x_+ = \max(x, 0)$ and $G(x)$ is the distribution function of the generalized extreme value distribution (GEV(μ, σ, ξ))

Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) \xrightarrow{d} \mathsf{G}(x)$$

then G must be the same type of the following form:

$$\mathsf{G}(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]_+^{\frac{-1}{\xi}}\right\}$$

where $x_+ = \max(x, 0)$ and $G(x)$ is the distribution function of the generalized extreme value distribution (GEV(μ, σ, ξ))

- μ and σ are location and scale parameters

Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) \xrightarrow{d} \mathsf{G}(x)$$

then G must be the same type of the following form:

$$\mathsf{G}(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]_+^{\frac{-1}{\xi}}\right\}$$

where $x_+ = \max(x, 0)$ and $G(x)$ is the distribution function of the generalized extreme value distribution (GEV(μ, σ, ξ))

- μ and σ are location and scale parameters
- ξ is a shape parameter determining the rate of tail decay, with

Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) \xrightarrow{d} \mathsf{G}(x)$$

then G must be the same type of the following form:

$$\mathsf{G}(x; \mu, \sigma, \xi) = \exp\left\{ -\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]_+^{\frac{-1}{\xi}} \right\}$$

where $x_+ = \max(x, 0)$ and $G(x)$ is the distribution function of the generalized extreme value distribution (GEV(μ, σ, ξ))

- μ and σ are location and scale parameters
- ξ is a shape parameter determining the rate of tail decay, with
    - $\xi > 0$ giving the heavy-tailed case (Fréchet)

Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \xrightarrow{d} \mathsf{G}(x)$$

then G must be the same type of the following form:

$$\mathsf{G}(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]_+^{\frac{-1}{\xi}}\right\}$$

where $x_+ = \max(x, 0)$ and $G(x)$ is the distribution function of the generalized extreme value distribution (GEV(μ, σ, ξ))

- μ and σ are location and scale parameters
- ξ is a shape parameter determining the rate of tail decay, with
  - $\xi > 0$ giving the heavy-tailed case (Fréchet)
  - $\xi = 0$ giving the light-tailed case (Gumbel)

Extremal Types Theorem (Fisher–Tippett 1928, Gnedenko 1943)

Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \xrightarrow{d} \mathsf{G}(x)$$

then G must be the same type of the following form:

$$\mathsf{G}(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]_+^{\frac{-1}{\xi}}\right\}$$

where $x_+ = \max(x, 0)$ and $G(x)$ is the distribution function of the generalized extreme value distribution (GEV(μ, σ, ξ))

- μ and σ are location and scale parameters
- ξ is a shape parameter determining the rate of tail decay, with
  - $\xi > 0$ giving the heavy-tailed case (Fréchet)
  - $\xi = 0$ giving the light-tailed case (Gumbel)
  - $\xi < 0$ giving the bounded-tailed case (reversed Weibull)

Example: Exponential Maxima

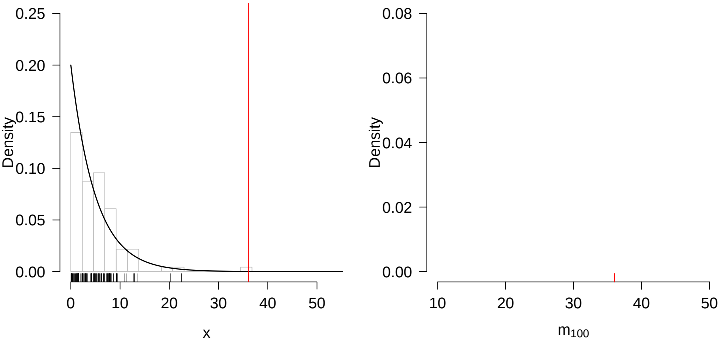
Let $X \sim \text{Exp}(\lambda = 1)$. Set $a_n = 1$, $b_n = \log(n)$. We want to show $\frac{M_n - b_n}{a_n}$ converges to a GEV distribution, where $M_n = \max_{i=1}^n X_i$.

$$\begin{aligned}
\mathbb{P}(\frac{M_n - b_n}{a_n} \leq x) &= \mathbb{P}(M_n \leq a_n x + b_n) \\
&= \mathbb{P}(M_n \leq x + \log(n)) \\
&= (1 - \exp(-x - \log(n)))^n \\
&= (1 - \frac{1}{n}\exp(-x))^n \\
&\xrightarrow{n\to\infty} \exp(-\exp(x))
\end{aligned}$$

It is the cdf of the standard Gumbel distribution

Extremal Types Theorem in Action

1. Generate 100 (n) random numbers from an Exponential distribution (population distribution)
2. Compute the sample maximum of these 100 random numbers
3. Repeat this process 120 times



Max-Stability and GEV

Definition

A distribution G is said to be max-stable if

$$G^k(a_k x + b_k) = G(x), \qquad k \in \mathbb{N}$$

for some constants $a_k > 0$ and b_k

- Taking powers of a distribution function results only in a change of location and scale
- A distribution is max-stable $\iff$ it is a GEV distribution

Quantiles and Return Levels

- Quantiles of GEV

$$G(x_p) = \exp\left\{-\left[1+\xi\left(\frac{x_p-\mu}{\sigma}\right)\right]_+^{\frac{-1}{\xi}}\right\} = 1-p$$

$$\Rightarrow x_p = \mu - \frac{\sigma}{\xi}\left[1-\{-\log(1-p)^{-\xi}\}\right] \qquad 0<p<1$$

- In the extreme value terminology, x_p is the return level associated with the return period $\frac{1}{p}$

Statistical Practice

Assume n is large enough so that

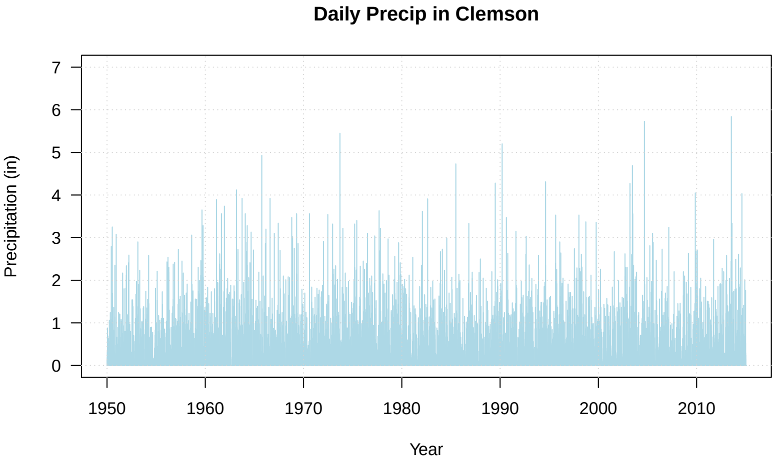
$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \approx \exp\left(-[1+\xi x]^{-1/\xi}\right)$$

$$\Rightarrow \mathbb{P}(M_n \le y) \approx \exp\left(-\left[1+\xi\left(\frac{y-b_n}{a_n}\right)^{-1/\xi}\right]\right)$$

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

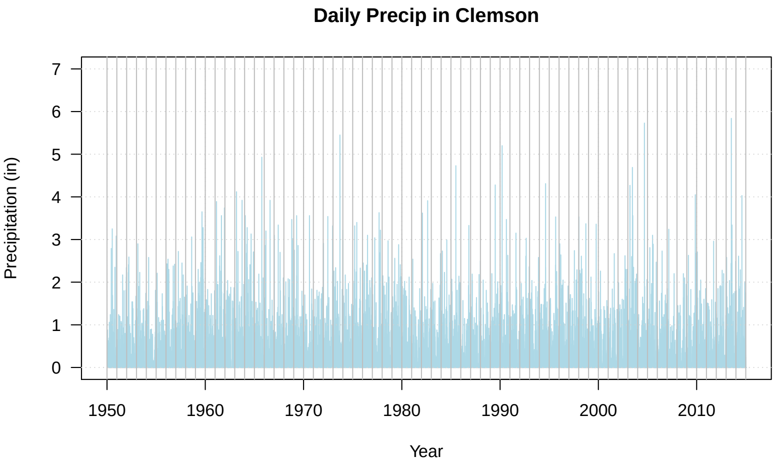
Then, we have a three-parameter estimation problem. μ, σ, ξ can be estimated via maximum likelihood

Clemson Daily Precipitation [Data Source: USHCN]



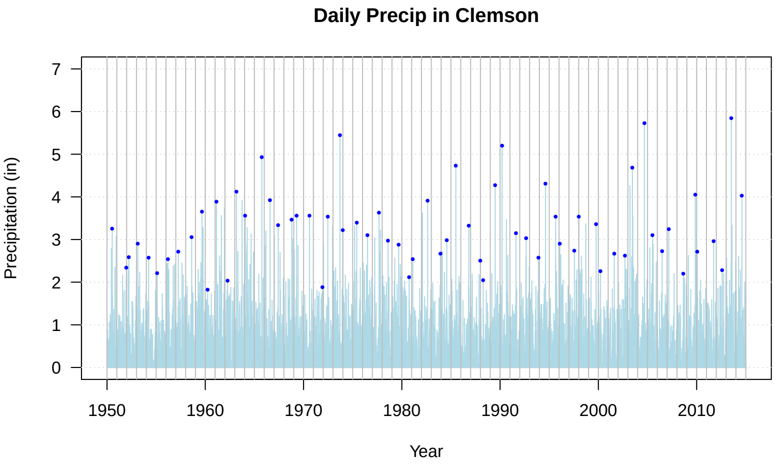
Block Maxima Method (Gumbel 1958)

1. Determine the block size and extract the block maxima



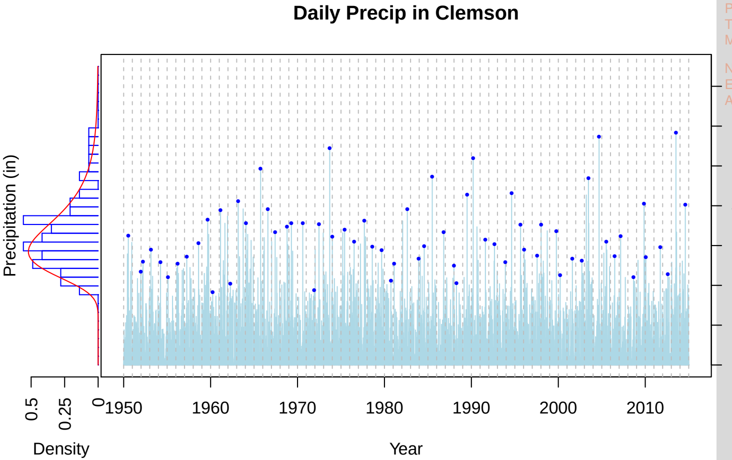
Block Maxima Method (Gumbel 1958)

1. Determine the block size and extract the block maxima



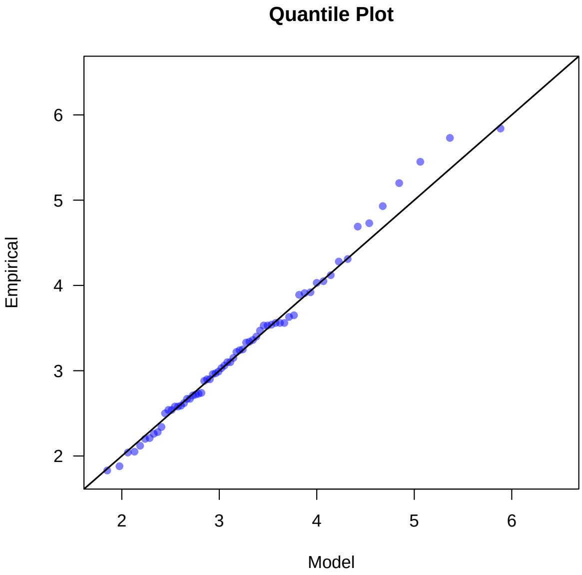
Block Maxima Method (Gumbel 1958)

2. Fit the GEV to the maximal and assess the fit



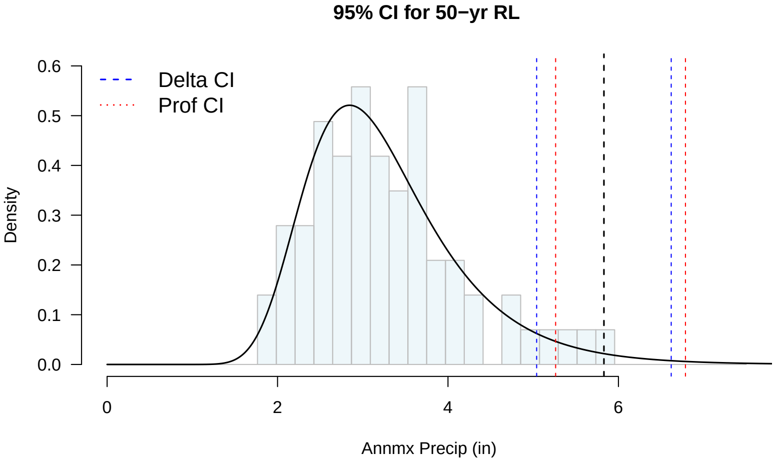
Block Maxima Method (Gumbel 1958)

2. Fit the GEV to the maximal and assess the fit



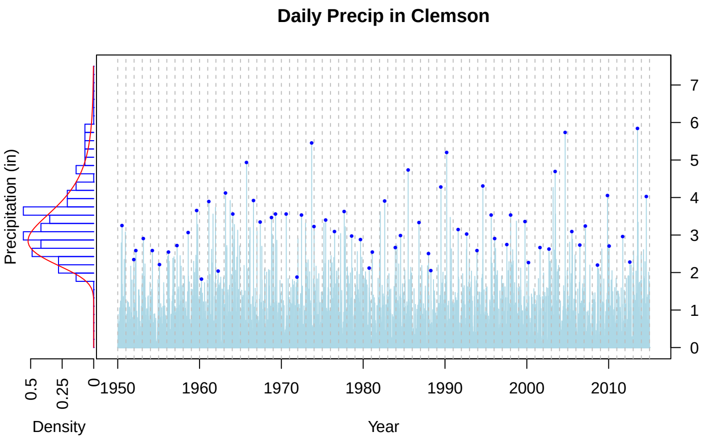
Block Maxima Method (Gumbel 1958)

3. Perform inference for return levels, probabilities, etc. Two methods to quantify estimation uncertainty: 1) Delta-method (relies on asymptotic normality); 2) Profile likelihood



Outline

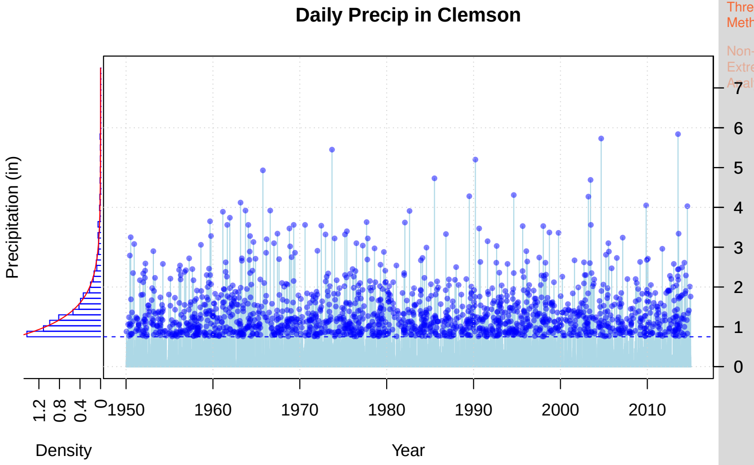
Recall the Block Maxima Method



Question: Can we use data more efficiently?

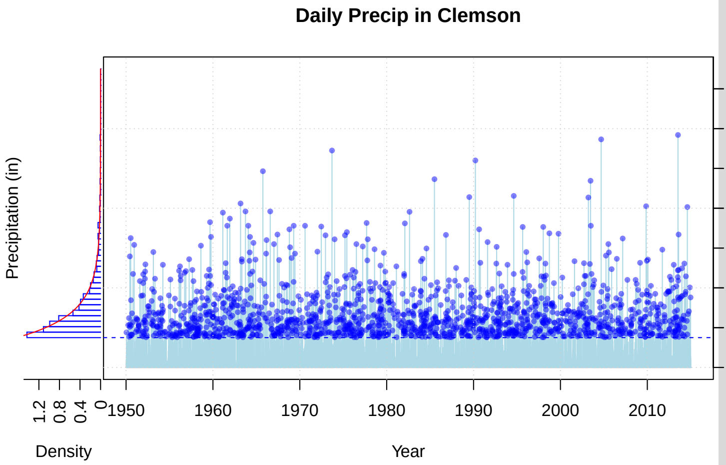
Peaks–over–threshold (POT) method [Davison & Smith 1990]

1. Select a "sufficiently large" threshold u, extract the exceedances



Peaks–over–threshold (POT) method [Davison & Smith 1990]

2. Fit an appropriate model to exceedances



GPD for Exceedances

If $M_n = \max_{i=1,\cdots,n} X_i$ (for a large n) can be apprximated by a GEV(μ, σ, ξ), then for sufficently large u,

$$\begin{aligned}
\mathbb{P}(X_i > x + u | X_i > u) &= \frac{n\mathbb{P}(X_i > x + u)}{n\mathbb{P}(X_i > u)} \\
&\to \left(\frac{1 + \xi \frac{x+u-b_n}{a_n}}{1 + \xi \frac{u-b_n}{a_n}} \right)^{\frac{-1}{\xi}} \\
&= \left(1 + \frac{\xi x}{a_n + \xi(u - b_n)} \right)^{\frac{-1}{\xi}}
\end{aligned}$$

$\Rightarrow$ Survival function of generalized Pareto distribution

Pickands–Balkema–de Haan Theorem (1974, 1975)

If $M_n = \max_{1 \le i \le n}\{X_i\} \approx$ GEV(μ, σ, ξ), then, for a “large” u (i.e., $u \to x_F = \sup\{x : F(x) < 1\}$),

$$\mathbb{P}(X > u) \approx \frac{1}{n}\left[1 + \xi\left(\frac{u-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}$$

$F_u = \mathbb{P}(X - u < y | X > u)$ is well approximated by the generalized Pareto distribution (GPD). That is:

$$F_u(y) \xrightarrow{d} H_{\tilde{\sigma},\xi}(y) \qquad u \to x_F$$

where

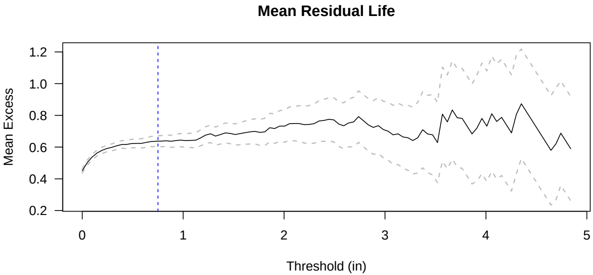
$$H_{\tilde{\sigma},\xi}(y) = \begin{cases} 1 - (1 + \xi y/\tilde{\sigma})^{-1/\xi} & \xi \neq 0; \\ 1 - \exp(-y/\tilde{\sigma}) & \xi = 0. \end{cases}$$

and $\tilde{\sigma} = \sigma + \xi(u - \mu)$

How to Choose the Threshold?

Bias–variance tradeoff:

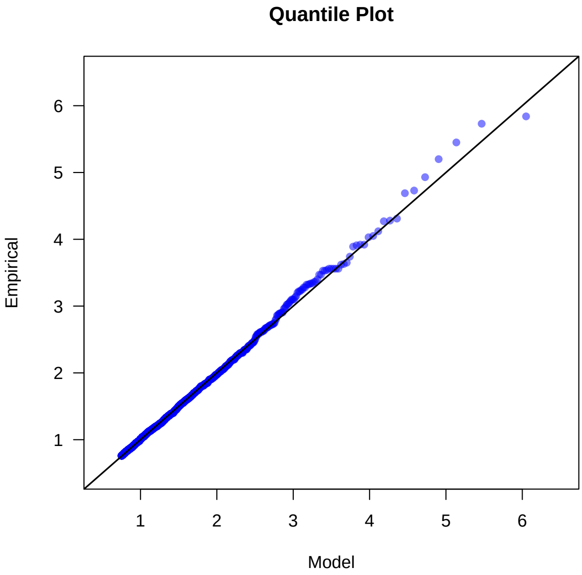
- Threshold too low ⇒ bias because of the model asymptotics being invalid
- Threshold too high ⇒ variance is large due to few data points



Task: To choose a u_0 s.t. the Mean Residual Life curve behaves linearly $\forall u > u_0$

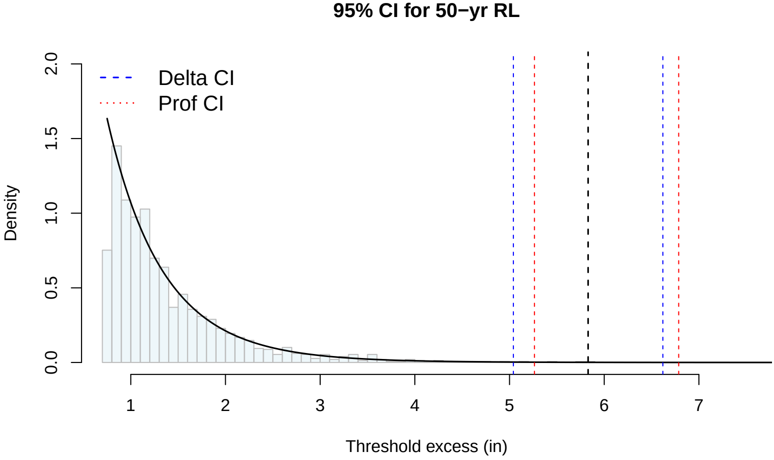
Peaks–over–threshold (POT) method [Davison & Smith 1990]

2. Fit an appropriate model to exceedances and assess the fit



Peaks–over–threshold (POT) method [Davison & Smith 1990]

3. Perform inference for return levels, probabilities, etc



Temporal Dependence

Question: Is the GEV still the limiting distribution for block maxima of a stationary (but not independent) sequence $\{X_i\}$?

Answer: Yes, as long as mixing conditions hold. (Leadbetter et al., 1983)

What does this mean for inference?

Block maximum approach: GEV still correct for marginal. Since block maximum data likely have negligible dependence, proceed as usual

Threshold exceedance approach: GPD is correct for the marginal. If extremes occur in clusters, estimation affected as likelihood assumes independence of threshold exceedances

Remarks on Univariate Extremes

- To estimate the tail, EVT uses only extreme observations
- Shape parameter ξ is extremely important but hard to estimate
- Threshold exceedance approaches allow the user to retain more data than block-maximum approaches, thereby reducing the uncertainty with parameter estimates
- Temporal dependence in the data is more of an issue in threshold exceedance models. One can either decluster, or alternatively, adjust inference

Outline

Modeling Non-Stationary GEV

- $M_t \sim \mathrm{GEV}(\mu(t), \sigma(t), \xi(t))$
- Typically assume “simple” structure for $\mu(t)$ and $\sigma(t)$, and $\xi(t)$ be a constant
- $\mu(t)$ and $\sigma(t)$ could depend on some relevant factors

Summary & Discussion

- Extreme value theory provides a framework to model extreme values
  - GEV for fitting block maxima
  - GPD for fitting threshold exceedances
  - Return level for communicating risk
- Practical Issues:

Summary & Discussion

- Extreme value theory provides a framework to model extreme values
  - GEV for fitting block maxima
  - GPD for fitting threshold exceedances
  - Return level for communicating risk
- Practical Issues: seasonality, temporal dependence, non-stationarity, ...

For Further Reading

- J. Beirlant, Y Goegebeur, J. Segers, and J Teugels
  Statistics of Extremes: Theory and Applications.
  Wiley, 2004.
- L. de Haan, and A. Ferreira
  Extreme Value Theory: An Introduction.
  Springer, 2006.
- S. I. Resnick
  Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer, 2007.