Lecture 11

Extreme Value Analysis

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

MATH 8090 Time Series Analysis
Week 11

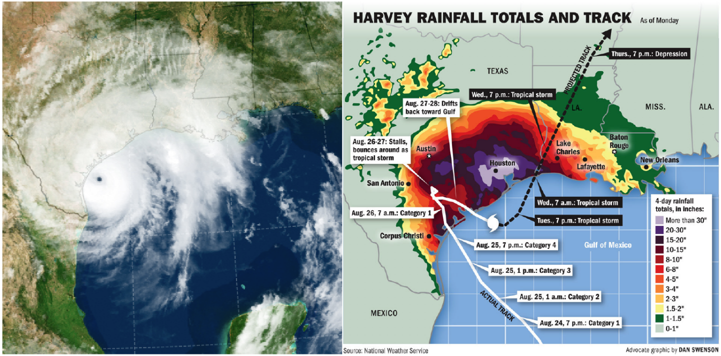
Whitney Huang
Clemson University

Agenda

1. **Motivation**
2. **Extremal Types Theorem & Block Maxima Method**
3. **Peaks–Over–Threshold (POT) Method**

Outline

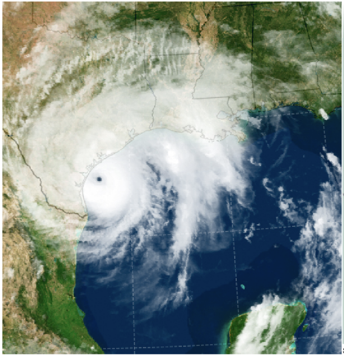
Extreme Rainfall During Hurricane Harvey

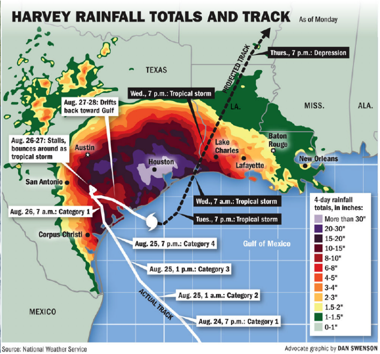


Source: NASA (Left); National Weather Service (Right)

- The highest total rainfall was 60.58 inches near Nederland, TX.

Extreme Rainfall During Hurricane Harvey

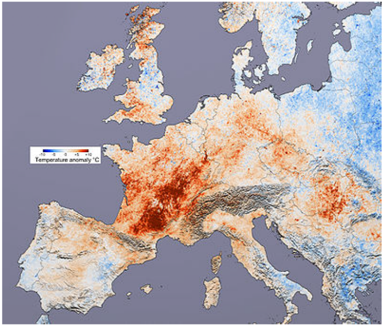




Source: NASA (Left); National Weather Service (Right)

- The highest total rainfall was 60.58 inches near Nederland, TX.
- **Annual average** rainfall for Nederland, TX: 59 inches

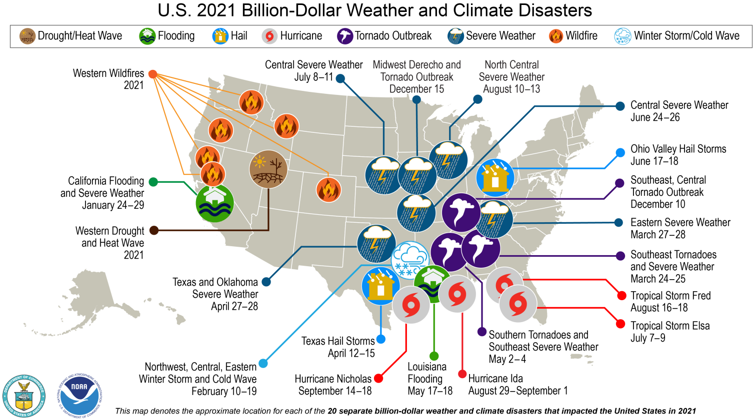
Environmental Extremes: Heatwaves, Storm Surges, etc.



- **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.

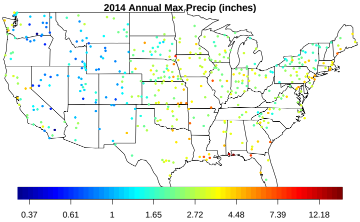


Source: National Oceanic and Atmospheric Administration

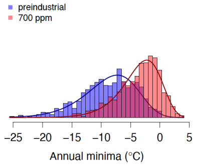
Goal: to quantify the tail behavior ⇒ often requires extrapolation.

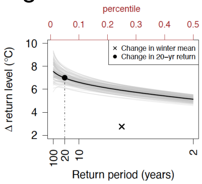
Some Scientific Questions

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



- How extremes change in future climate conditions?

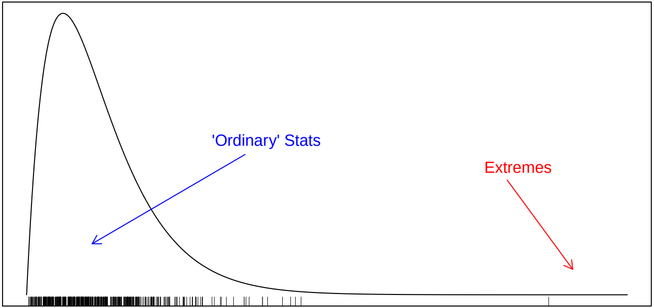




Outline

1. Motivation
2. **Extremal Types Theorem & Block Maxima Method**
3. Peaks–Over–Threshold (POT) Method

Usual vs Extremes



	Target	Theory	Distribution
Ordinary Stats	bulk distribution	CLT	Normal
Extreme Stats	tail distribution(s)	?	?

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

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

⇒ 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 $S_n = \sum_{i=1}^{n} X_i$

⇒ 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

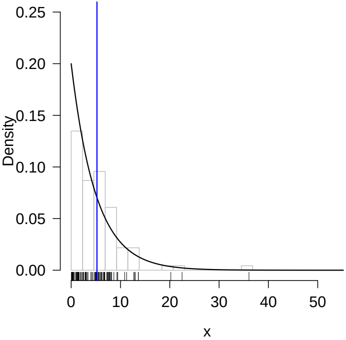
Question: Can we get the limiting distribution of

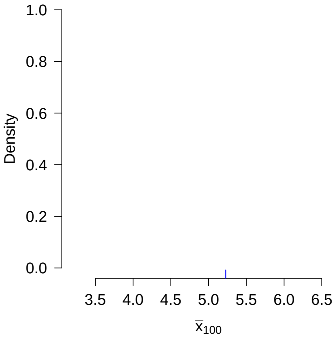
$$\frac{M_n - b_n}{a_n}$$

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

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) \overset{d}{\to} \mathrm{G}(x)$$

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

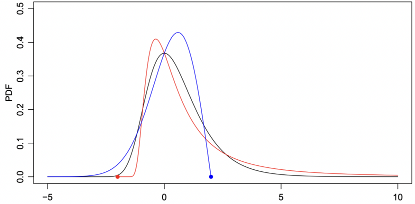
Define $M_n = \max\{X_1, \cdots, X_n\}$ where $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$. If $\exists\, a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$, if

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) \overset{d}{\to} \mathrm{G}(x)$$

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

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

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



- $\xi > 0$: Fréchet (heavy-tail)
- $\xi = 0$: Gumbel (light-tail)
- $\xi < 0$: reversed Weibull (short-tail)

Example: Exponential Maxima

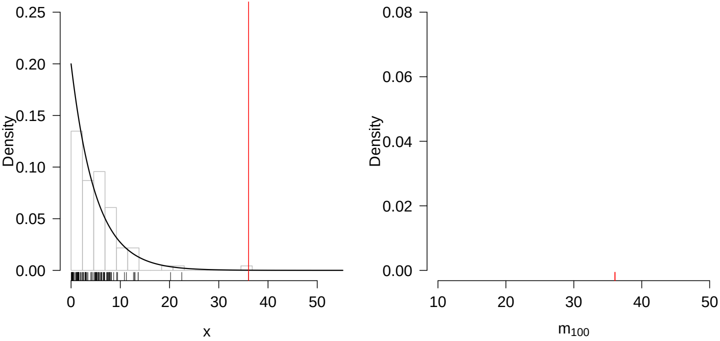
Let $X \sim \text{Exp}(\lambda = 1)$. Set $a_n = 1$, $b_n = \log(n)$. We want to show $\frac{M_n - b_n}{a_n}$ converges to a GEV distribution, where $M_n = \max_{i=1}^n X_i$.

$$\begin{aligned}
\mathbb{P}\left(\frac{M_n - b_n}{a_n} \le x\right) &= \mathbb{P}(M_n \le a_n x + b_n) \\
&= \mathbb{P}(M_n \le x + \log(n)) \\
&= (1 - \exp(-x - \log(n)))^n \\
&= \left(1 - \frac{1}{n}\exp(-x)\right)^n \\
&\overset{n\to\infty}{\longrightarrow} \exp(-\exp(x))
\end{aligned}$$

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

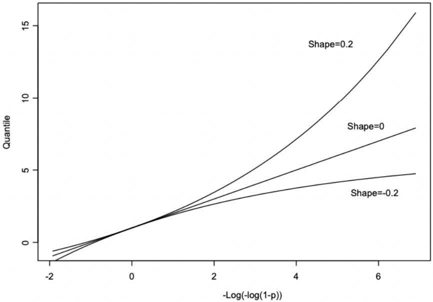
- Taking powers of a distribution function results only in a change of location and scale
- A distribution is max-stable $\iff$ it is a GEV distribution

Quantiles and Return Levels

- Quantiles of GEV

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

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



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

Statistical Practice

Assume n is large enough so that

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

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

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

Then, we have a three-parameter estimation problem. μ, σ, ξ can be estimated via maximum likelihood[1]

[1] Probability weighted moments/L-moments and Bayesian methods can also be used to carry out parameter estimation

Maximum Likelihood Estimation

Let $M_1, \cdots, M_k \overset{\text{iid}}{\sim} \text{GEV}$, then log-likelihood for (μ, σ, ξ) when $\xi \neq 0$ is

$$\ell(\mu, \sigma, \xi) = -k \log \sigma - (1 + 1/\xi) \sum_{i=1}^{k} \log \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^{k} \left[1 + \xi \left(\frac{m_i - \mu}{\sigma} \right) \right]^{-1/\xi},$$

provided that $1 + \xi(\frac{m_i - \mu}{\sigma}) > 0$, for $i = 1, \cdots, k$.

When $\xi = 0 \rightarrow$ use the Gumbel limit of the GEV

Maximum likelihood estimate (MLE) is obtained by (numerically) maximization of log-likelihood shown above

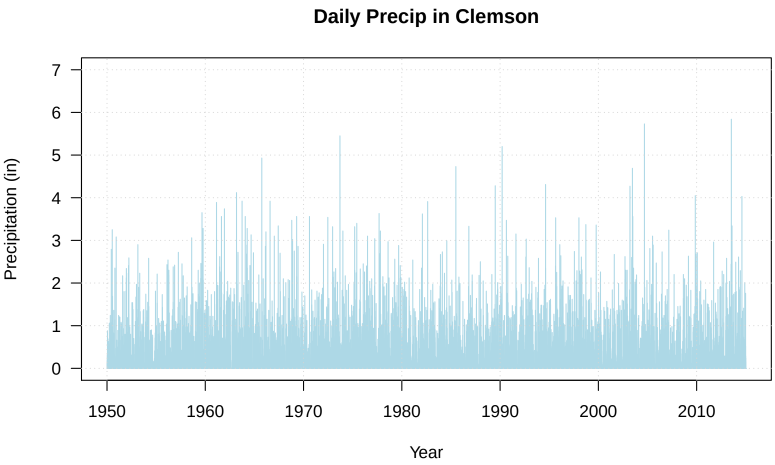
Uncertainty Quantification for GEV Estimation

Parameters not very interpretable. Better to provide uncertainty about a more meaningful quantity (e.g. 100-year return level)

Two methods:

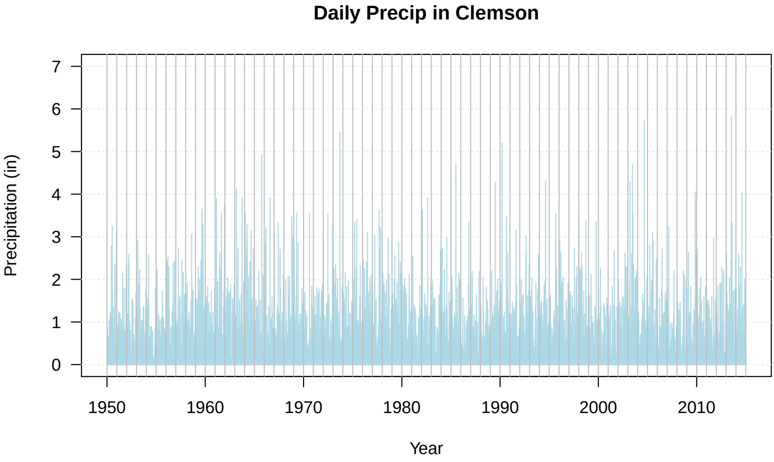
- Delta method
  - +: easy to compute with a closed form expression
  - -: symmetric confidence interval is not realistic (especially for long return levels)
- Profile likelihood method
  - +: can allow for asymmetric confidence intervals
  - -: need to compute numerically

Clemson Daily Precipitation [Data Source: USHCN]



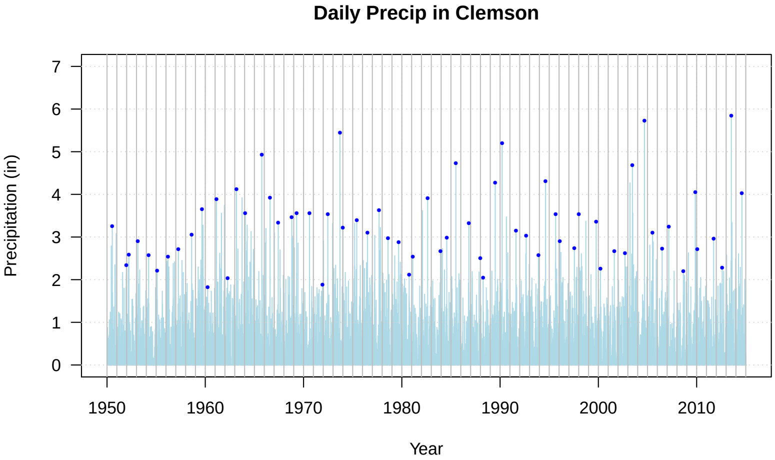
Block Maxima Method (Gumbel 1958)

1. Determine the block size and extract the block maxima



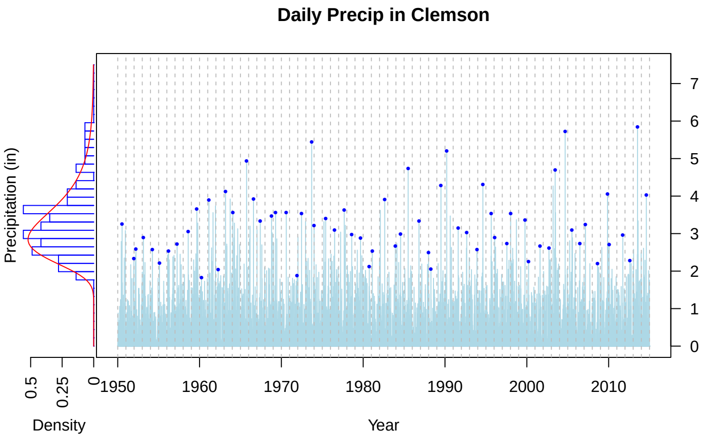
Block Maxima Method (Gumbel 1958)

1. Determine the block size and extract the block maxima



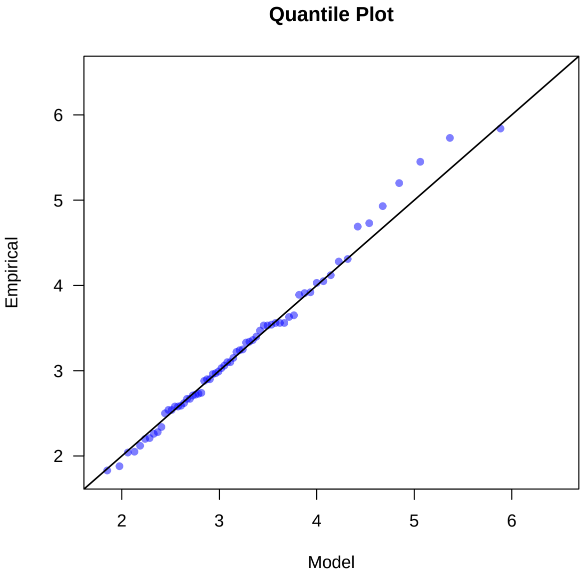
Block Maxima Method (Gumbel 1958)

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



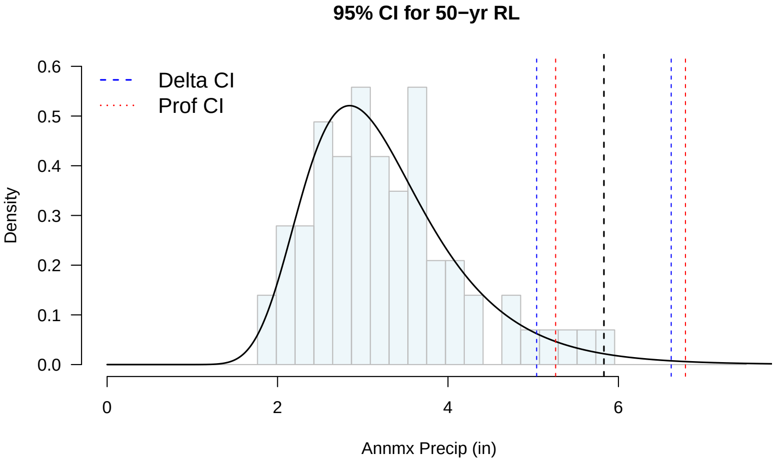
Block Maxima Method (Gumbel 1958)

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



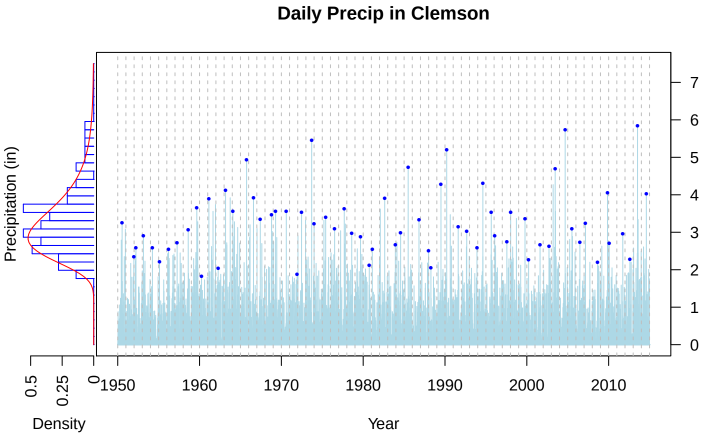
Block Maxima Method (Gumbel 1958)

3. Perform inference for return levels, probabilities, etc.



Outline

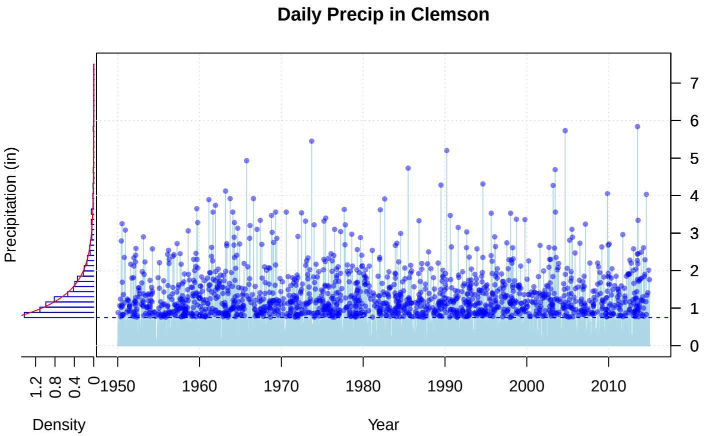
Recall the Block Maxima Method



Question: Can we use data more efficiently?

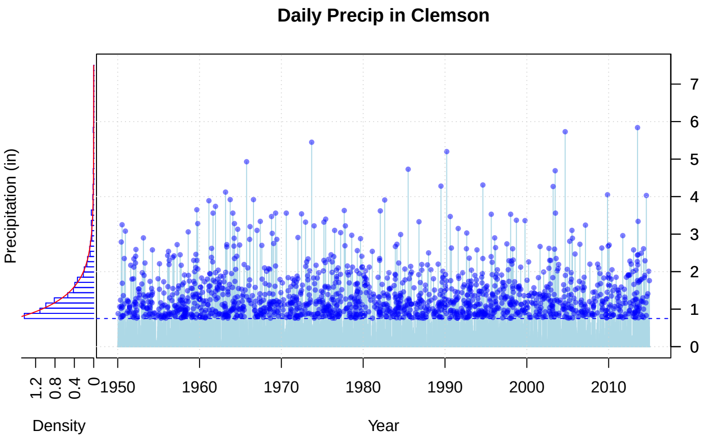
Peaks–over–threshold (POT) method [Davison & Smith 1990]

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,

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

⇒ Survival function of generalized Pareto distribution

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

If $M_n = \max_{1\leq i\leq n}\{X_i\} \approx$ 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

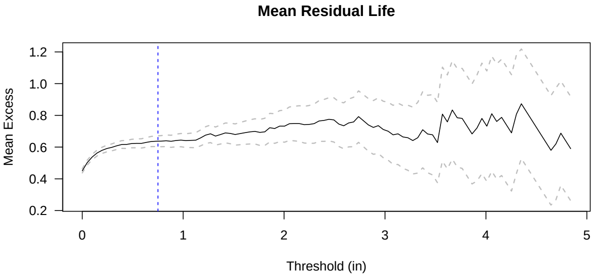
$$H_{\tilde{\sigma},\xi}(y) = \begin{cases} 1 - (1 + \xi y/\tilde{\sigma})^{-1/\xi} & \xi \neq 0; \\ 1 - \exp(-y/\tilde{\sigma}) & \xi = 0. \end{cases}$$

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

How to Choose the Threshold?

Bias–variance tradeoff:

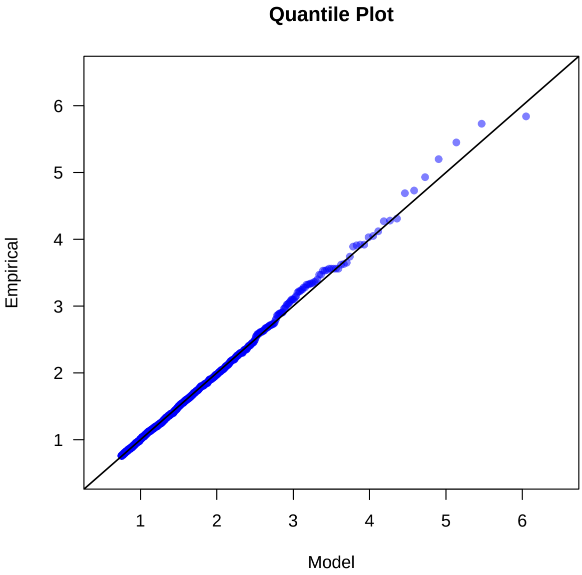
- Threshold too low ⇒ bias because of the model asymptotics being invalid
- Threshold too high ⇒ variance is large due to few data points



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

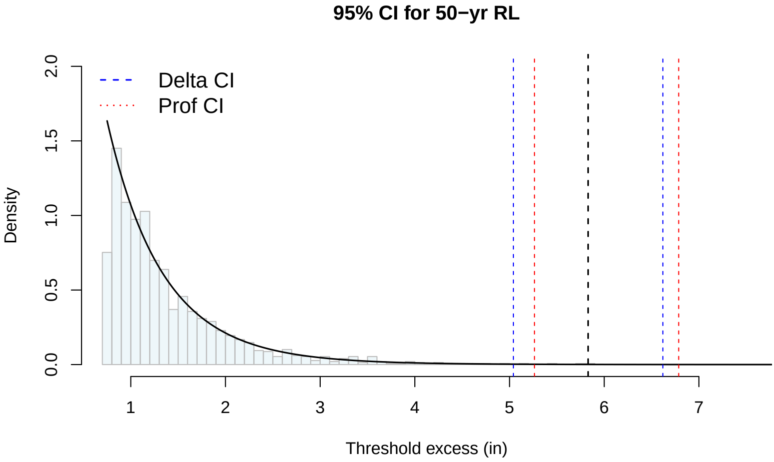
Peaks–over–threshold (POT) method [Davison & Smith 1990]

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)

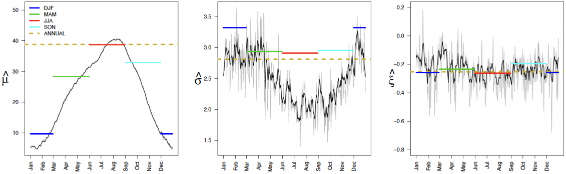
What does this mean for inference?

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

Modeling Non-Stationary Extremes: Seasonality and Long-Term Trend

- $M_t \sim \text{GEV}(\mu(t), \sigma(t), \xi(t))$



- Typically assume fairly simple structure for $\mu(t)$ and $\sigma(t)$,

  e.g. $\mu(t) = \mu_0 + \mu_1 t,$

  and $\xi(t)$ be a constant

- $\mu(t)$ and $\sigma(t)$ could depend on some **physically-informed factors** (e.g. Clausius-Clapeyron precipitation-temperature scaling)

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

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.
- Dipak Dey and Jun Yan
  Extreme Value Modeling and Risk Analysis: Methods and Applications. CRC Press, 2016.