EVA Tutorial #3

ISSUES ARISING IN EXTREME VALUE ANALYSIS

Rick Katz

Institute for Mathematics Applied to Geosciences
National Center for Atmospheric Research
Boulder, CO USA

email: rwk@ucar.edu
Home page: www.isse.ucar.edu/staff/katz/
Lecture: www.isse.ucar.edu/extremevalues/docs/eva3.pdf
Outline

(1) Penultimate Approximations
(2) Origin of Bounded and Heavy Tails
(3) Clustering at High Levels
(4) Complex Extreme Events
(5) Risk Communication under Stationarity
(6) Risk Communication under Nonstationarity
(1) Penultimate Approximations

- “Ultimate” Extreme Value Theory
  -- GEV distribution as limiting distribution of maxima
  \[ X_1, X_2, \ldots, X_n \text{ independent with common cdf } F \]
  \[ M_n = \max \{ X_1, X_2, \ldots, X_n \} \]

- Penultimate Extreme Value Theory
  -- Suppose \( F \) in domain of attraction of Gumbel type (i.e., \( \xi = 0 \))
  -- Still preferable in nearly all cases to use GEV as approximate
distribution for maxima (i.e., act as if \( \xi \neq 0 \))
-- Expression (as function of block size $n$) for shape parameter $\xi_n$

“Hazard rate” (or “failure rate”):

$$H_F(x) = F'(x) / [1 - F(x)]$$

Instantaneous rate of “failure” given “survived” until $x$

Alternative expression: $H_F(x) = -[\ln(1 - F)]'(x)$

One choice of shape parameter (block size $n$):

$$\xi_n = (1/H_F)'(x) \mid_{x=u(n)}$$

Here $u(n)$ is “characteristic largest value”

$$u(n) = F^{-1}(1 - 1/n)$$

[or $(1 - 1/n)$th quantile of $F$]
-- Because $F$ assumed in domain of attraction of Gumbel,

$$\xi_n \to 0 \text{ as block size } n \to \infty$$

-- More generally, can use behavior of $H_F(x)$ for large $x$ to determine domain of attraction of $F$

In particular, if

$$(1/H_F)'(x) \to 0 \text{ as } x \to \infty$$

then $F$ is in domain of attraction of Gumbel

*Note:* Straightforward to show that hazard rate of lognormal distribution satisfies above condition (i.e., in domain of attraction of Gumbel)
• Example: Exponential Distribution

-- Exact exponential upper tail (unit scale parameter)

\[ 1 - F(x) = \exp(-x), \quad x > 0 \]

-- Penultimate approximation

Hazard rate: \[ H_F(x) = 1, \quad x > 0 \]

(Constant hazard rate consistent with memoryless property)

Shape parameter: \[ \xi_n = 0 \]

So no benefit to penultimate approximation
• *Example:* Normal Distribution (with zero mean & unit variance)

-- Fisher & Tippett (1928) proposed Weibull type of GEV as penultimate approximation

Hazard rate: \( H_\Phi(x) \approx x, \) for large \( x \)

[Recall that \( 1 - \Phi(x) \approx \varphi(x) / x \)]

Characteristic largest value: \( u(n) \approx (2 \ln n)^{1/2}, \) for large \( n \)

Penultimate approximation is Weibull type with

\[
\xi_n \approx -1 / (2 \ln n)
\]

For example: \( \xi_{100} \approx -0.11, \xi_{365} \approx -0.085 \)
• **Example**: “Stretched Exponential” Distribution

-- Traditional form of Weibull distribution (Bounded below)

\[ 1 - F(x) = \exp(-x^c), \quad x > 0, \quad c > 0 \]

where \( c \) is shape parameter (unit scale parameter)

Hazard rate: \( H_F(x) = c x^{c-1}, \quad x > 0 \)

Characteristic largest value: \( u(n) = (\ln n)^{1/c} \)

Penultimate approximation has shape parameter

\[ \xi_n \approx (1 - c) / (c \ln n) \]

(i) \( c > 1 \) implies \( \xi_n \uparrow 0 \) as \( n \to \infty \) (i. e., Weibull type)

(ii) \( c < 1 \) implies \( \xi_n \downarrow 0 \) as \( n \to \infty \) (i. e., Fréchet type)
(2) Origin of Bounded and Heavy Tails

- Upper Bounds / Penultimate approximation

-- Weibull type of GEV (i.e., \(\xi < 0\))

For instance, provides better approximation than Gumbel type when “parent” distribution \(F\):

(i) Normal (e.g., for temperature)

(ii) Stretched exponential with \(c > 1\) (e.g., for wind speed)

-- Apparent upper bound

Complicates interpretation (e.g., “thermostat hypothesis” or maximum intensity of hurricanes)
• Heavy tails / Penultimate approximation

-- Fréchet type of GEV (i. e., $\xi > 0$)

For instance, provides better approximation than Gumbel when parent distribution $F$:

Stretched exponential distribution with $c < 1$

-- Possible explanation for apparent heavy tail of precipitation

Wilson & Toumi (2005):

Based on physical argument, proposed stretched exponential with $c = 2/3$ (Universal value, independent of season or location) as distribution for heavy precipitation
-- Simulation experiment

Generated observations from stretched exponential distribution with shape parameter $c = 2/3$

Determine maximum of sequence of length $n = 100$, $M_{100}$
(Annual maxima: Daily precipitation occurrence rate $\approx 27\%$)

Annual prec. maxima: Typical estimated $\xi \approx 0.10$ to $0.15$
(Penultimate approximation gives $\xi_{100} \approx 0.11$)

Fitted GEV distribution (Sample size = 1000):

  Obtained estimate of $\xi \approx 0.12$
Q-Q Plot: Stretched exponential simulation
• Heavy Tails / Chance mechanism

-- Mixture of exponential distributions

Suppose $X$ has exponential distribution with scale parameter $\sigma^*$:

$$\Pr\{X > x \mid \sigma^*\} = \exp\left[-\left(\frac{x}{\sigma^*}\right)\right], \quad x > 0, \quad \sigma^* > 0$$

Further assume that the rate parameter $\nu = 1/\sigma^*$ varies according to a gamma distribution with shape parameter $\alpha$ (unit scale), pdf:

$$f_\nu(\nu; \alpha) = \left[\Gamma(\alpha)\right]^{-1} \nu^{\alpha-1} \exp(-\nu), \quad \alpha > 0$$

The unconditional distribution of $Y$ is heavy-tailed:

$$\Pr\{X > x\} = (1 + x)^{-\alpha}$$

(i.e., exact GP distribution with shape parameter $\xi = 1/\alpha$)
-- Simulation experiment

Induce heavy tail from conditional light tails

Let rate parameter of exponential distribution have gamma distribution with shape parameter $\alpha = 2$

Then unconditional (mixture) distribution is GP with shape parameter $\xi = 0.5$

Fit GP distribution to simulated exponential mixture (Sample size = 1000):

   Obtained estimate of $\xi \approx 0.51$
(3) Clustering at High Levels

• As example, consider stationary Gaussian process
  
  -- Joint distribution of $X_t$ and $X_{t+k}$ is bivariate normal with
    autocorrelation coefficient $\rho_k$, $k = 1, 2, \ldots$
  
  -- So consider two random variables $(X, Y)$ with bivariate normal
    distribution with correlation coefficient $\rho$, $|\rho| < 1$

  No “clustering at high levels” (in asymptotic sense; i.e., extremal
  index $\theta = 1$):

  $$\Pr\{Y > u \mid X > u\} \rightarrow 0 \text{ as } u \rightarrow \infty$$
Bivariate normal distribution (\(\rho = 0.75\))

Simulation (sample size = 10,000)
Bivariate normal distribution (\(\rho = 0.75\))

Simulation (sample size = 10,000)
Bivariate normal distribution

$\Pr(Y > u \mid X > u)$

Threshold $u$

- $\rho = 0$
- $\rho = 0.25$
- $\rho = 0.5$
- $\rho = 0.75$
• Interpretation of extremal index $\theta$, $0 < \theta \leq 1$

(i) Mean cluster length $\approx 1/\theta$

(ii) Effective sample size
(as if take maximum of $n^* = n\theta$ “unclustered” observations)

*Note:* Does not resemble same concept based on time averages

Effect of $\theta < 1$ on GEV distribution:

Adjustment to location and scale parameters, $\mu$ and $\sigma$, but no adjustment to shape parameter $\xi$

In block maxima approach, effect of $\theta < 1$ automatically subsumed in fitted parameters of GEV (could affect approximation accuracy)
“Intervals estimator” of extremal index $\theta$ (Ferro-Segers 2003)

-- “Interexceedance” times (i.e., time between exceedances)

(i) If $X_t > u$ & $X_{t+1} > u$, then interexceedance time = 1

(ii) If $X_t > u$, $X_{t+1} < u$, $X_{t+2} > u$, then interexceedance time = 2, etc.

Coefficient of variation (i.e., st. dev. / mean) of interexceedance times converges to function of $\theta$ as threshold $u \to \infty$

Does not require identification of clusters (could chose runs declustering parameter $r$ so that mean cluster length $\approx 1/\theta$)

-- Confidence interval for $\theta$

Resample interexceedance times (because of extremal dependence, need to modify conventional bootstrap)
A Gaussian first-order autoregressive process with $\rho_1 = 0.25$.
Gaussian first-order autoregressive process with $\rho_1 = 0.75$
Evidence of clustering at high levels

Fort Collins summer maximum temperature

Extremal index

Threshold u (°F)

Evidence of clustering at high levels
Lack of evidence of clustering at high levels
(4) Complex Extreme Events

- Heat waves
  -- Extreme weather phenomenon
  -- Lack of use of statistical methods based on extreme value theory
  -- Complex phenomenon / Ambiguous concept
  -- Focus on hot spells instead
    (Derive more full-fledged heat waves from model for hot spells)
  -- Devise simple model (only use univariate extreme value theory)
  -- Simple enough to incorporate trends (or other covariates)
• Start with point process (or Poisson-GP) model

-- Rate of occurrence of clusters
   Modeled as Poisson process (rate parameter $\lambda$)

-- Intensity of cluster
   Cluster maxima modeled as GP distribution (shape parameter $\xi$, scale parameter $\sigma^*$)

• Retain clusters ("hot spells"), rather than declustering

-- Model cluster statistics
   (i) Duration (e. g., geometric distribution with mean $1/\theta$)
   (ii) Dependence of excesses within cluster (conditional GP model)
• Model for excesses with cluster (runs parameter $r = 1$)

Let $Y_1, Y_2, \ldots, Y_k$ denote excesses over threshold within given cluster / spell (assume of length $k > 1$)

(i) Model first excess $Y_1$ as unconditional GP distribution (instead of cluster maxima)

(ii) Model conditional distribution of $Y_2$ given $Y_1$ as GP with scale parameter depending on $Y_1$; e.g., with linear link function

$$
\sigma^*(y) = \sigma_0^* + \sigma_1^* y, \text{ given } Y_1 = y
$$

Similar model for conditional distribution of $Y_3$ given $Y_2$ (etc.)

Requires only univariate extreme value theory (not multivariate)
Phoenix Maximum Temperature

- Median
- Lower quartile
- Upper quartile

Second exceedance (°F)

First exceedance (°F)
• Conditional distribution of $Y_2$ given $Y_1 = y$

-- Conditional mean  [increases with $\sigma^*(y)$]

$$E(Y_2 \mid Y_1 = y) = \sigma^*(y) / (1 - \xi), \; \xi < 1$$

-- Conditional variance  (increases with mean)

$$\text{Var}(Y_2 \mid Y_1 = y) = [E(Y_2 \mid Y_1 = y)]^2 / (1 - 2 \xi), \; \xi < 1/2$$

-- Conditional quantile function

$$F^{-1}[\rho; \sigma^*(y), \xi] = [\sigma^*(y) / \xi] [(1 - \rho)^{-\xi} - 1], \; 0 < \rho < 1$$

Increases more rapidly with $\sigma^*(y)$ for higher $\rho$
• Introduction of trends

-- Cluster rate
  Trend in mean of Poisson rate parameter $\lambda(s)$, year $s$

-- Cluster length
  Trend in mean of geometric distribution $1/\theta(s)$, year $s$

-- Cluster maxima (or first excess)
  Trend in scale parameter of GP distribution $\sigma^*(s)$, year $s$

-- Other covariates such as index of atmospheric blocking
Phoenix (GLM with log link, $P$-value $\approx 0.01$)
(5) Risk Communication under Stationarity

- Interpretation of return level $x(p)$ (under stationarity)

-- Stationarity implies identical distributions
  (not necessarily independence)

(i) Expected waiting time (under temporal independence)

Waiting time $W$ has geometric distribution:

$$\Pr\{W = k\} = (1 - p)^{k-1} p, \ k = 1, 2, \ldots, \ E(W) = 1/p$$

(ii) Length of time $T_p$ for which expected number of events = 1

$$1 = \text{Expected no. events} = T_p p, \ \text{so} \ T_p = 1/p$$
(6) Risk Communication under Nonstationarity

- Options

-- Retain one of these two interpretations

  Not clear which one is preferable:
  Property (ii) is easier to work with (like average probability)
  Property (i) may be more meaningful for risk analysis

-- Switch to “effective” return period and “effective” return level
  (i. e., quantiles varying over time)
Moving flood plain from year-to-year (not necessarily feasible?)
- Alternative concept

-- Extreme event \( X_t > u \)

-- Choose threshold \( u \) to achieve desired value of

\[
\text{Pr\{One or more events over time interval of length } T\}
\]

-- Under stationarity (and temporal independence)

As an example, if \( p = 0.01 \) (i.e., 100-yr return level):

\[
\text{Pr\{one or more events over 30 yrs\}} = 1 - (0.99)^{30} \approx 0.26
\]

\[
\text{Pr\{one or more events over 100 yrs\}} = 1 - (0.99)^{100} \approx 0.63
\]