3 Statistics of Extremes slide 67

3.1 Introduction slide 68

Basic ideas

- Section 2.3 described models for extremes in which **background data** $x_1, \ldots, x_{mt_0}$ are treated as a realisation of $X_1, \ldots, X_{mt_0} \stackrel{\text{iid}}{\sim} F$.
- This is highly idealised, since in applications
  - the models are asymptotic, but the data are finite, so there may be bias;
  - data are very often not identically distributed, owing to seasonality, trend or dependence on external factors;
  - data are typically dependent, owing to short-term persistence of extreme conditions;
  - there may be other complications, e.g., selection of data because they are extreme or missing data or ...
- Despite this the **extremal paradigm**, i.e., fitting asymptotic models to finite-sample data, is widely used, and is the basis of extremal analysis.

Minima

- In general discussion we consider maxima and large values — what about minima and small values?
- As
$$Y = \min(X_1, \ldots, X_m) = -\max(-X_1, \ldots, -X_m) = -Y^-,$$
say, we see that if we apply the arguments of §2.3 to $-X$, then
$$\tilde{G}(y) \approx \mathrm{P}(Y \leq y) = \mathrm{P}(Y^- \geq -y) \approx 1 - G(-y),$$
where G is the GEV approximation for $\max(-X_1, \ldots, -X_m)$. Hence
$$\tilde{G}(y; \tilde{\eta}, \tilde{\tau}, \tilde{\xi}) = 1 - G(-y; -\eta, \tau, \xi),$$
where G is estimated from the negative minima.

Estimation

- ☐ Mostly we use maximum likelihood estimation according to the recipe on slide 24.
- ☐ This has theoretical and practical advantages:
  - it is efficient (has the smallest possible variance) in large samples — in regular situations (more later);
  - likelihood ratio tests are generally fairly powerful;
  - there's a simple recipe to follow — write down the likelihood and maximise it — which works in many situations;
  - lots of code already exists and can be readily applied. Hooray!
- ☐ Other methods of estimation are also used:
  - method of moments estimation to get initial values for maximising a likelihood;
  - probability-weighted (or L-) moments estimation is widely used in hydrology and some other domains, because it can beat ML estimation in small samples;
  - in more complex problems the likelihood can be awkward, and then other methods must be used.

Moment estimation

- ☐ Define moments for random variable X as $\mu'_r = \mathrm{E}(X^r)$ for $r = 1, \ldots$ (if μ'_r finite).
- ☐ If X depends on $p \times 1$ parameter vector θ, then $\mu'_r = \mu'_r(\theta)$, and we estimate θ by solving the equations

$$\mu'_r(\theta) = n^{-1} \sum_j X_j^r, \quad r = 1, \ldots, p.$$

- ☐ Moment estimators usually simple but inefficient (variance larger than for competing approaches)
- ☐ For GEV, μ'_r exists only if $\xi r < 1$, so must have $\xi < 1/3$ to estimate all three parameters, and $\xi < 1/6$ for them to have finite variances. Much too restrictive for use in practice.
- ☐ Useful for finding starting-values for ML estimation.

L-moment estimation

- Define **probability-weighted moments** as $\mu'_{r,s,t} = \mathrm{E}[X^r F(X)^s \{1 - F(X)\}^t]$ for $r, s, t = 0, 1, 2, \ldots$, or equivalently

$$\mu'_{r,s,t} = \int_0^1 x_p^r p^s (1-p)^t dp, \quad \text{where} \quad F(x_p) = p;$$

  ordinary moments have $s = t = 0$.
- Use $\beta_s \equiv \mu'_{1,s,0}$ for $s = 0, 1, \ldots$ to fit GEV and GPD.
- In practice estimate the **L-moments**, $\lambda_1 = \beta_0$, $\lambda_2 = 2\beta_1 - \beta_0$, $\ldots$, by

$$\widehat{\lambda}_1 = \frac{1}{\binom{n}{1}} \sum_{j=1}^n X_{(j)}, \quad \widehat{\lambda}_2 = \frac{1}{2\binom{n}{2}} \sum_{j=1}^n \left\{ \binom{j-1}{1} - \binom{n-j}{1} \right\} X_{(j)}, \quad \ldots,$$

- L-moment estimators of η, τ and ξ based on $\widehat{\lambda}_1$, $\widehat{\lambda}_2$ and $\widehat{\lambda}_3$ are linear in the observations, so are more robust than the ordinary moment estimators.
- Have good small-sample properties, but don't generalise to complex settings.

Plan

- Now discuss
  - basic models,
  - exploratory methods,
  - fitting and interpretation and
  - model checking

  for basic models for maxima and for threshold exceedances.
- Then discuss targets of inference — measures of risk — and practical complications.

Extremal Types Theorem

Theorem 15 (Extremal types) *Let $M = \max(X_1, \ldots, X_m)$ be the maximum of a random sample $X_1, \ldots, X_m$. If sequences of real numbers $\{a_m\} > 0$ and $\{b_m\}$ can be chosen so that the centred and scaled sample maximum, $Y_m = (M - b_m)/a_m$, has a non-degenerate limiting distribution G, then this must be the* generalized extreme-value distribution (GEV),

$$G(y) = \begin{cases} \exp\left[-\{1+\xi(y-\eta)/\tau\}_+^{-1/\xi}\right], & \xi \neq 0, \\ \exp\left[-\exp\{-(y-\eta)/\tau\}\right], & \xi = 0, \end{cases} \quad y \in \mathbb{R}, \tag{11}$$

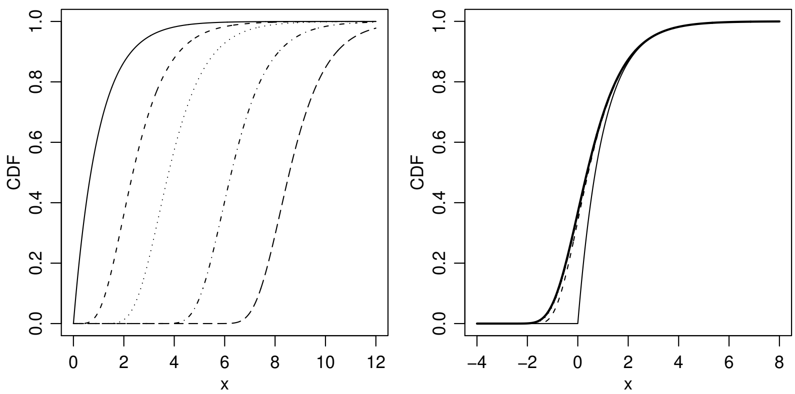
where $a_+ = \max(a, 0)$ for any real a, and with $\xi, \eta \in \mathbb{R}$ and $\tau > 0$. Put another way, $Y_m \xrightarrow{D} Y \sim G$ as $m \to \infty$, giving the 'finite-m' approximation $\mathrm{P}(Y_m \le y) \approx G(y)$.

- ☐ The 'types', which arise for $\xi = 0$, $\xi > 0$ and $\xi < 0$, are now usually subsumed into (11), and are discussed below.
- ☐ This theorem provides a single distribution for maxima, and is in some ways stronger than the Central Limit Theorem, since we only assume that linear rescaling can result in a non-degenerate distribution, without other assumptions on F.
- ☐ This is a natural model for maxima (and minima by using $1 - G(-y)$).

http://stat.epfl.ch slide 76

Examples

Example 16 *Find sequences $\{a_m\}$ and $\{b_m\}$ such that maxima of independent variables from the (a) uniform, (b) exponential, and (c) Pareto distributions have non-degenerate limiting distributions.*



Distributions of maxima (left) and renormalized maxima (right) of $m = 1$, 7, 30, 365, 3650 standard exponential variables (from left to right), with limiting Gumbel distribution (heavy).

http://stat.epfl.ch slide 77

Note I to Example 16

- ☐ Note that

$$\mathrm{P}\left\{(M-b_m)/a_m \le y\right\} = \mathrm{P}\left\{M \le b_m + a_m y\right\} = F^m(b_m + a_m y),$$

and we need to choose a_m and b_m such that this has a limit as $m \to \infty$. We saw from Theorem 14 that a limit $G(y) = \exp\{-\Lambda(y)\}$, so it is equivalent to identify Λ.

- ☐ (a) In the uniform case, $F(x) = x$ for $x \in [0,1]$. Provided $0 \le b_m + a_m y \le 1$, we therefore have

$$F(b_m + a_m y)^m = (b_m + a_m y)^m,$$

so if we set $b_m = 1$, $a_m = 1/m$ and $-m \le y \le 0$, we have $(b_m + a_m y)^m \to e^y$. Hence

$$\Lambda(y) = \begin{cases} -y, & y \le 0, \\ 0, & y > 0, \end{cases}$$

i.e., $\Lambda(y) = (-y)_+$. Clearly Λ is decreasing on $(-\infty, 0)$. Hence

$$G(y) = \exp\{-\Lambda(y)\} = \begin{cases} e^y, & y \le 0, \\ 1, & y > 0, \end{cases}$$

which is the distribution function of $-W$, where $W \sim \exp(1)$. It is straightforward to check that this G is (11) with $\eta = 1$, $\tau = 1$ and $\xi = -1$.

- ☐ (b) In the exponential case, $F(x) = 1 - \exp(-x)$ for $x > 0$. Provided $b_m + a_m y > 0$,

$$F(b_m + a_m y)^m = \left[1 - \exp\left\{-(b_m + a_m y)\right\}\right]^m,$$

so if we set $b_m = \log m$ and $a_m = 1$, and if $y > -\log m$,

$$G(y) = \lim_{m\to\infty} F(b_m + a_m y)^m = \lim_{m\to\infty}\left(1 - \frac{e^{-y}}{m}\right)^m = \exp\left(-e^{-y}\right), \quad y \in \mathbb{R},$$

which is (11) with $\eta = 0$, $\tau = 1$ and $\xi = 0$. Here $\Lambda(y) = e^{-y}$ with support in $\mathbb{R}$.

- ☐ (c) In the Pareto case, $F(x) = 1 - x^{-\alpha}$ for $x > 1$ and $\alpha > 0$. Provided $b_m + a_m y > 1$, we have

$$F(b_m + a_m y)^m = \left\{1 - (b_m + a_m y)^{-\alpha}\right\}^m$$

so if we set $b_m = 0$ and $a_m = m^{1/\alpha}$, and if $y > m^{-1/\alpha}$, we have

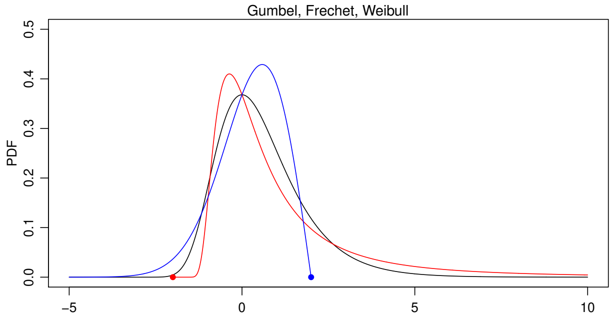
$$G(y) = \lim_{m\to\infty} F(b_m + a_m y)^m = \lim_{m\to\infty}\left(1 - \frac{y^{-\alpha}}{m}\right)^m = \exp\left(-y^{-\alpha}\right), \quad y \ge 0,$$

which is (11) with $\eta = 1$, $\tau = 1/\alpha$ and $\xi = 1/\alpha$. In this case

$$\Lambda(y) = \begin{cases} \infty, & y \le 0, \\ y^{-\alpha}, & y > 0. \end{cases}$$

- ☐ Note that we have not shown that the three limits above are the only ones possible, just that we can choose a_m and b_m to obtain these limits.

GEV and ‘three types’



- ξ is a shape parameter determining the rate of tail decay, with:
  - $\xi > 0$ giving the heavy-tailed **Fréchet (Type II)** density with support $(\eta - \tau/\xi, \infty)$;
  - $\xi = 0$ giving the light-tailed **Gumbel (Type I)** density, with support $\mathbb{R}$;
  - $\xi < 0$ giving the short-tailed **(reverse) Weibull (Type III)** density, with support $(-\infty, \eta - \tau/\xi)$.
- The usual Weibull distribution gives a model for minima.
- η and τ are location and scale parameters (not so crucial as the shape parameter ξ).

Properties of the GEV

- **Support**: If $\xi > 0$ then $Y > \eta - \tau/\xi$, and if $\xi < 0$ then $Y < \eta - \tau/\xi$.
- **Moments**: $\mathrm{E}(Y^r)$ exists only if $\xi < 1/r$, so the mean exists only if $\xi < 1$, the variance only if $\xi < 1/2$, etc. In applications (particularly in finance) some moments may not exist.
- **Quantiles**: solve $G(y) = p$ for $0 < p < 1$, but usually we use the **return levels** given by solving $G(y_p) = 1 - p$ (next slide) — so y_p is the $(1 - p)$ quantile (careful!)
- **Maximum likelihood estimation**: is regular only if $\xi > -1/2$. Not usually a problem in applications.
- **Max-stability**: if $Y_1, \ldots, Y_T \overset{\text{iid}}{\sim}$ GEV(η, τ, ξ) then $\max(Y_1, \ldots, Y_T) \sim$ GEV(η_T, τ_T, ξ_T), i.e.,

$$G(y; \eta, \tau, \xi)^T = G(y; \eta_T, \tau_T, \xi_T)$$

  where

$$\eta_T = \begin{cases} \eta + \tau(T^\xi - 1)/\xi, & \xi \neq 0, \\ \eta + \tau \log T, & \xi = 0, \end{cases} \qquad \tau_T = \tau T^\xi, \quad \xi_T = \xi,$$

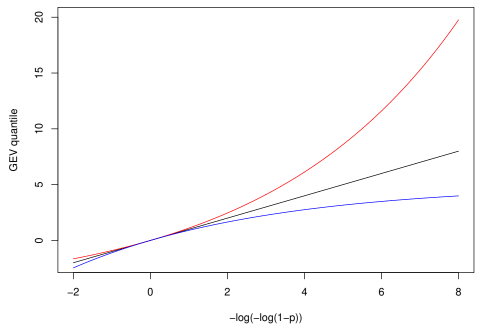
  so the distribution type and shape parameter are unchanged by taking maxima.
- In fact the GEV is the only max-stable class of distributions.

Quantiles and return levels

- Define the **return level** associated to the **return period** $T = 1/p$ (blocks) as

$$y_p = \eta + \tau \frac{\{-\log(1-p)\}^{-\xi} - 1}{\xi}, \quad 0 < p < 1,$$

  i.e., the solution to $G(y_p) = 1 - p = 1 - 1/T$.
- Informally, y_p is the level expected to be exceeded once every T blocks.
- The plot below compares the quantiles for $\xi = -0.2$ (blue) and $\xi = 0.2$ (red) with the Gumbel quantiles (black).



Statistical approach

- Assume background data $x_1, x_2, \ldots$ are IID realisations from some continuous distribution F to which the GEV approximation applies.
- Take maxima $y = \max(x_1, \ldots, x_m)$ of blocks of size m from the background data.
  - for environmental time series, typically $m \approx 365$ for annual maxima, $m \approx 30$ for monthly maxima, ...
  - in finance, typically $m = 250$ for annual maxima, $m = 20$ for monthly maxima, ...
- Suppose the resulting series of maxima $y_1, \ldots, y_n$ are IID GEV(η, τ, ξ).
- Fit the GEV by maximum likelihood estimation and use the fitted model for inferences.

Exploratory plot for maxima

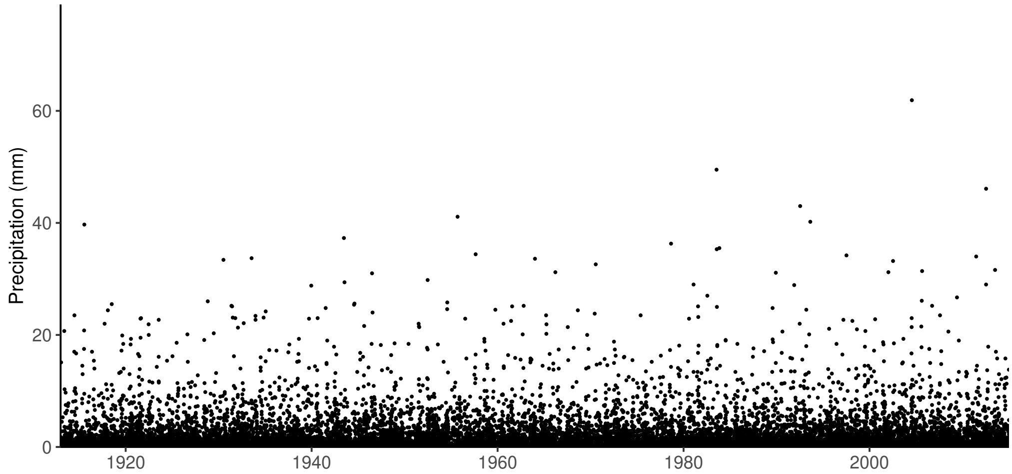
- Plot ordered block maxima $y_{(1)} \leq \cdots \leq y_{(n)}$ against Gumbel plotting positions

$$-\log\left[-\log\left\{j/(n+1)\right\}\right], \quad j = 1, \ldots, n.$$

- After allowing for noise,
  - convex shape suggests $\xi > 0$,
  - straight line suggests $\xi \approx 0$,
  - concave shape suggests $\xi < 0$.
- Outliers, heavy rounding or other issues with data should be visible.
- Comparison of these plots for different block sizes may also suggest a minimum block size for the GEV to apply.

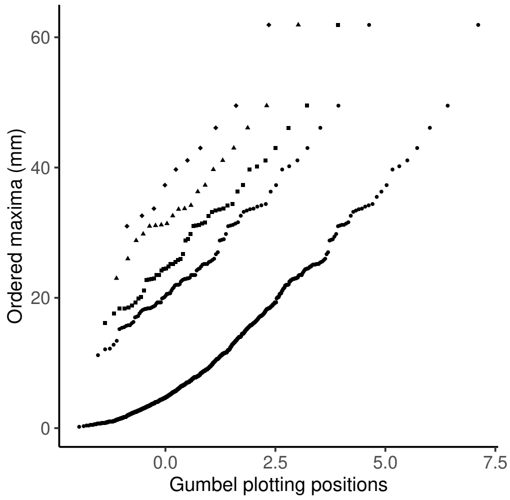
Abisko daily rainfall data

- Daily precipitation in Abisko, in northern Sweden, 1913–2014. The largest value is 61.9 mm, but many values are zero and most of the positive values are quite small.



Abisko block maxima

- Gumbel QQplot of maxima for blocks of lengths (from bottom) one month and one, two, five and ten years.



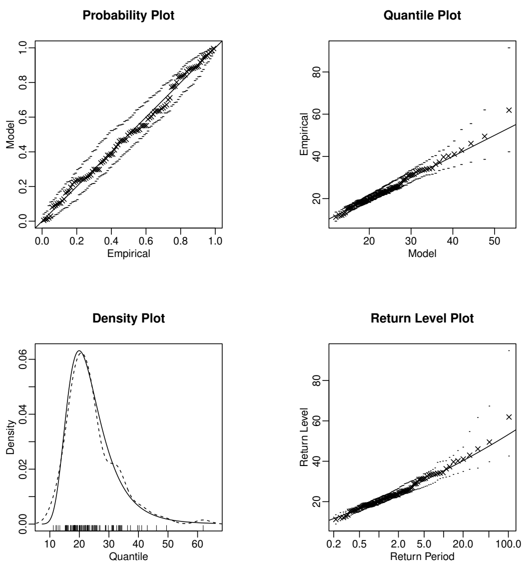
Abisko annual maxima

- ☐ QQplot suggests stability from one year onwards, with slight convexity, so let's fit the GEV to annual maxima:

```
library(evd)
(fit <- fgev(year.max))

Call: fgev(x = year.max)
Deviance: 691.9509

Estimates
     loc     scale     shape
20.40530   5.84596   0.08353

Standard Errors
    loc    scale    shape
0.64854  0.48317  0.07193

Optimization Information
  Convergence: successful
  Function Evaluations: 27
  Gradient Evaluations: 7
```


Abisko annual maxima

- ☐ Let's check the fit using `plot(fit)`:



Commentary

- ☐ These (horrible!) plots use the fitted GEV CDF $\widehat{G} \equiv G(\cdot; \widehat{\eta}, \widehat{\tau}, \widehat{\xi})$ and are the
  - **probability plot** showing $\left\{(j/(n+1), \widehat{G}(y_{(j)})) : j = 1, \ldots, n\right\}$, which should be a straight line of unit gradient if $\widehat{G}$ is a good fit;
  - **quantile plot** showing $\left\{(\widehat{G}^{-1}\{j/(n+1)\}, y_{(j)}) : j = 1, \ldots, n\right\}$, which should be a straight line of unit gradient if $\widehat{G}$ is a good fit;
  - **return level plot** showing (solid line) $(-\log(1-p), \widehat{G}^{-1}(1-p))$, for $0 < p < 1$, and the points $\{(-\log\{j/(n+1)\}, y_{(j)}) : j = 1, \ldots, n\}$, which should lie on the line if $\widehat{G}$ is a good fit;
  - **density plot** showing a kernel density estimate based on $y_1, \ldots, y_n$ (shown by the rug) and the fitted GEV density.
- ☐ Some of the plots have pointwise 95% limits for individual points.
- ☐ They show essentially the same information but on different scales to highlight different aspects of the fit.
- ☐ In this case the fit seems reasonable.

3.3 Basic Methods for Exceedances

Basic ideas

- ☐ Background data $x_1, \ldots, x_{mt_0}$ comprise t_0 blocks each of m observations.
- ☐ Model the exceedances over some threshold u by a Poisson process with measure

$$\mu\{(t', t) \times [x, \infty)\} = (t - t')\Lambda(x), \quad 0 \leq t' < t \leq t_0, \quad x > u,$$

where

$$\Lambda(x) = \left(1 + \xi\frac{x - \eta}{\tau}\right)_+^{-1/\xi}.$$

- ☐ This implies that the times of exceedances are a Poisson process of rate $p_u = \Lambda(u)$ in $(0, t_0)$ and the exceedance sizes are IID with GP distribution

$$\mathrm{P}(X_j - u \leq x \mid X_j > u) = 1 - (1 + \xi x/\sigma_u)_+^{-1/\xi},$$

where $\sigma_u = \tau + \xi(u - \eta)$.

- ☐ This yields two fitting approaches:
  - estimate η, τ and ξ directly by fitting the Poisson process likelihood;
  - estimate σ_u and ξ from the exceedances and p_u from the number of exceedances, n_u.
- ☐ The second, **peaks over thresholds (POT)**, approach is most used in practice, as it's easier to explain and understand, but both fits are equivalent.

Exceedance Theorem

Theorem 17 (Exceedance) *Let X be a random variable having distribution function F, and suppose that a function c_u can be chosen so that the limiting distribution of $(X-u)/c_u$, conditional on $X > u$, is non-degenerate as u approaches the upper support value $x^* = \sup\{x : F(x) < 1\}$ of X. If such a limiting distribution exists, it must be of* generalized Pareto *form, i.e.,*

$$H(x) = \begin{cases} 1 - (1+\xi x/\sigma)_+^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-x/\sigma), & \xi = 0, \end{cases} \quad x > 0, \tag{12}$$

where $\xi \in \mathbb{R}$ and $\sigma > 0$. Expression (12) *is the* **generalized Pareto distribution (GPD)**.

- ☐ There is a close connection with the extremal types theorem, which applies for maxima under the same conditions as the exceedance theorem applies for exceedances, and with the same ξ.
- ☐ The GPD is a natural model for exceedances over high thresholds (and under low ones, using $1 - H(-x)$).

Example 18 *Find a limiting distribution for threshold exceedances for $Z \sim N(0,1)$. Recall that $1-\Phi(z) \sim \phi(z)/z$ as $z \to \infty$.*

Note to Example 18

- ☐ Here $x^* = \infty$ and for large z we have $1 - \Phi(z) \sim \phi(z)/z$.
- ☐ By analogy with renormalising maxima we aim to find a function $c_u > 0$ such that

$$\lim_{u\to\infty} \mathrm{P}\{(Z-u)/c_u > x \mid Z > u\}$$

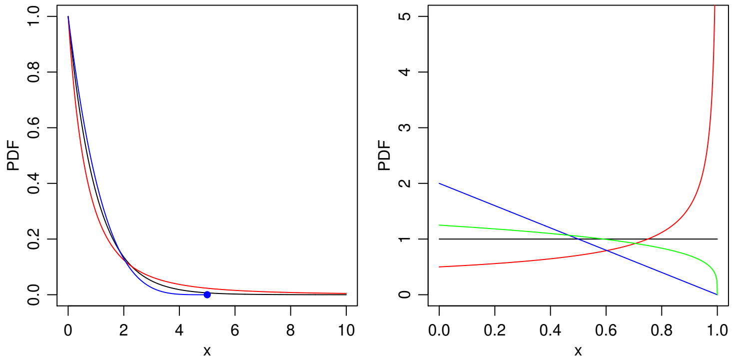
is non-degenerate. The hint gives that for fixed $x > 0$ and large u,

$$\begin{aligned} \mathrm{P}\{(Z-u)/c_u > x \mid Z > u\} &= \frac{\mathrm{P}(Z > u + c_u x)}{\mathrm{P}(Z > u)} \\ &= \frac{1 - \Phi(u + c_u x)}{1 - \Phi(u)} \\ &\sim \frac{\phi(u + c_u x)/(u + c_u x)}{\phi(u)/u} \\ &= \frac{u}{u + c_u x} \exp\{u^2/2 - (u + c_u x)^2/2\} \\ &= \frac{1}{1 + c_u x/u} \exp(-c_u u x - c_u^2 x^2/2), \end{aligned}$$

so if we choose $c_u = 1/u$ then the ratio tends to unity and the exponent tends to $-x$, i.e., the limiting distribution for an appropriately rescaled exceedance is standard uniform.

- ☐ If we had chosen $c_u = 1/(\sigma u)$ for any fixed $\sigma > 0$ we would have an exponential limit, with mean σ, as in (12), so we can think of the parameter σ as arising because we don't know the ideal scaling function.

Generalized Pareto distribution



- A flexible distribution whose density can take a variety of shapes.
- Left: exponential density ($\xi = 0$, black), heavy-tailed density ($\xi = 0.5$, red) and light-tailed density ($\xi = -0.2$, blue, with upper terminal shown); all have $\sigma = 1$.
- Right: densities with negative shape parameter and upper terminal at $x = 1$, with $\xi = -1$ (black), $\xi = -2$ (red), $\xi = -0.5$ (blue) and $\xi = -0.8$ (green).

Stability and threshold choice

- Both approaches require a threshold u to be chosen. Note that
  - the Poisson process parameters should be **stable** above an appropriate threshold u,
  - u too low will lead to bias (model inappropriate) and u too high will increase variance (too few exceedances).
- If the Poisson process model is stable above $u_{\min}$, then estimates of η, τ and ξ should be similar for $u > u_{\min}$, but will become more variable for higher u.
- If $X \sim \text{GPD}(\sigma, \xi)$, then $X - u \mid X > u \sim \text{GPD}(\sigma + \xi u, \xi)$, and this implies that

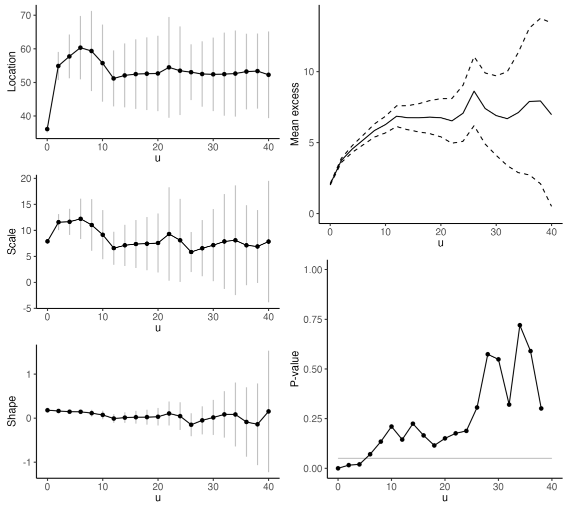
$$\text{E}(X - u \mid X > u) = \frac{\sigma + \xi u}{1 - \xi}, \quad \xi < 1,$$

so a **mean excess plot (or mean residual life plot)** of

$$\frac{\sum_j (x_j - u) I(x_j > u)}{\sum_j I(x_j > u)} \quad \text{against} \quad u,$$

should be approximately straight with slope $\xi/(1-\xi)$ above $u_{\min}$.
- Can also test for equal shape parameters above u (Northrop–Coleman test).

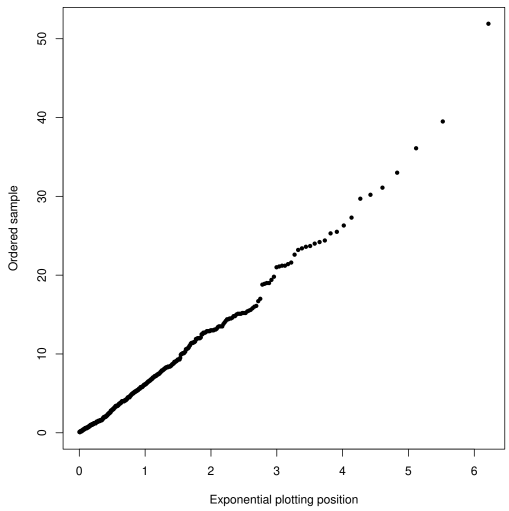
Abisko threshold analysis



☐ All panels suggest that $u_{\min}$ is reasonable.

Exploratory plot

☐ The natural plot here is of ordered exceedances against exponential plotting positions:

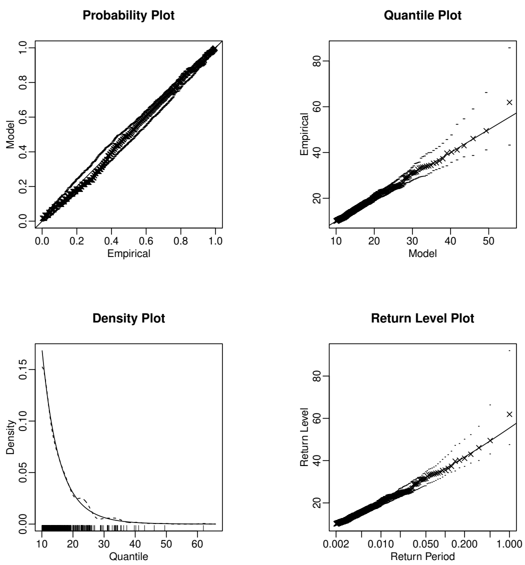


GPD fit

```
(fit.gpd <- fpot(abisko$precip,threshold=10))

Deviance: 2828.05

Threshold: 10
Number Above: 499
Proportion Above: 0.033

Estimates
  scale    shape
5.83261  0.07025

Standard Errors
  scale    shape
0.39483  0.05088

Optimization Information
  Convergence: successful
  Function Evaluations: 16
  Gradient Evaluations: 6
```


Abisko POT fit

☐ Let's check the fit using `plot(fit.gpd)`:

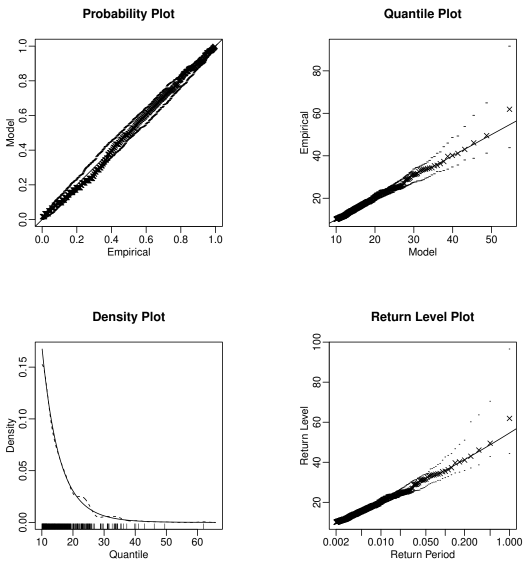


Poisson process fit

```
(fit.pp <- fpot(abisko$precip, threshold=10, model="pp", npp=365.25,
   start=list(loc=20,scale=6.5,shape=0.01)))
# needs initial values and number of points/block

Deviance: 2241.606

Threshold: 10
Number Above: 499
Proportion Above: 0.0134

Estimates
     loc     scale     shape
19.79658   6.52110   0.07026

Standard Errors
    loc    scale    shape
0.55597  0.37895  0.05088

Optimization Information
  Convergence: successful
  Function Evaluations: 20  ...  Gradient Evaluations: 8
```


Abisko Poisson process fit

☐ Let's check the fit using plot(fit.pp):



Summary

- The three fits agree fairly well:
  - Maxima: $\widehat{\eta} = 20.4_{0.649}$, $\widehat{\tau} = 5.84_{0.483}$, $\widehat{\xi} = 0.08_{0.072}$;
  - Poisson process: $\widehat{\eta} = 19.8_{0.556}$, $\widehat{\tau} = 6.52_{0.379}$, $\widehat{\xi} = 0.07_{0.051}$;
  - POT: $\widehat{p}_u = 0.033$, $\widehat{\sigma}_u = 5.83_{0.394}$, $\widehat{\xi} = 0.07_{0.051}$.
- The location and scale parameters are estimated quite well, but the shape much less well.
- The shape parameter estimate is slightly positive, but not significantly so (some hydrologists claim that rainfall has $\xi \approx 0.1 \ldots$).
- The fit appears to be good.
- In applications one would need to check that the threshold fits are robust to the choice of u (above $u_{\min}$).
- It is tempting to fit the model with $\xi = 0$, which will give much smaller standard errors for the other parameters. But as we do not know that $\xi = 0$, this reduction in uncertainty may be unrealistic, and it may introduce bias in extrapolation.

http://stat.epfl.ch slide 99

3.4 Targets of Inference slide 100

Return levels and return periods

- In basic analyses, typically aim to estimate risk measures such as

$$\mathrm{P}(X > x), \quad x_p = F_X^{-1}(1-p),$$

where X is a background observation and x and x_p are larger than any data,
  - e.g., legal requirement for nuclear installations to estimate the highest windspeed in $T = 10^7$ years, so if there are daily data, then $p = 1/(365.25T)$.
- x_p is a **T-year return level** with a **return period** of $1/p$ observations or T years.
- The return level solves the equation

$$F^{N_p}(x_p) = 1 - p,$$

where N_p is the number of background observations in the return period.

http://stat.epfl.ch slide 101

Return levels and return periods II

- ☐ Solving

$$F^{N_p}(x_p) = 1 - p$$

for the POT model gives

$$x_p = u + \frac{\sigma_u}{\xi}\left[\left\{\frac{1-(1-p)^{1/N_p}}{p_u}\right\}^{-\xi} - 1\right], \quad x_p > u, \tag{13}$$

where p_u is the probability that a single background observation exceeds u.

- ☐ The GEV applies to maxima of blocks of m background observations, so we effectively take

$$1 - p = G^{N_p/m}(x_p), \tag{14}$$

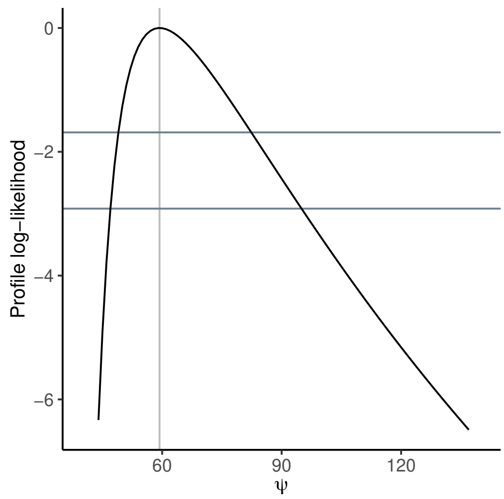
which yields

$$x_p = \mu + \frac{\sigma}{\xi}\left[\{-m\log(1-p)/N_p\}^{-\xi} - 1\right]. \tag{15}$$

- ☐ Both formulae are replaced by their limits as $\xi \to 0$ for the Gumbel or exponential fits.
- ☐ Point estimates of both are obtained by using the fitted parameter values.

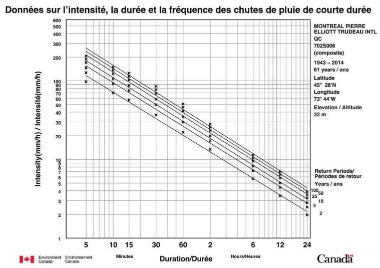
Profile log-likelihood

- ☐ Here ψ is the 100-year return value for daily precipitation at Abisko based on the GEV fit.
- ☐ The strong asymmetry means that symmetric confidence intervals could be very misleading.



Return levels and return periods III

- In hydrology, an **intensity-duration-frequency (IDF)** curve describes the relationship between rainfall intensity, duration, and a given return period and is used for flood risk assessment and water management.
- For each duration D, the frequency and magnitude of extreme rainfall events are estimated.
- Relying on the GEV applied to the series of annual maxima, estimates of x_p, the T-year return level, are produced. For comparison purposes, we work with $I = x_p/D$.
- The Gumbel distribution is usually used for convenience but more general approaches have recently been proposed.



IDF curves for Montréal airport. Source: Environment and Climate Change Canada (ECCC)

Other measures of risk

- In environmental applications it may be important to estimate amounts of rain falling into an entire catchment area, or the length and impact of a heatwave, or ...
- The Basel Accords regulate measures of risk to be used by financial institutions:
  - the **Value at Risk** VaR_p is another name for a quantile/return level x_p;
  - the **Expected Shortfall** is defined as the expected loss conditional on VaR_p being exceeded,

$$\text{E}(X - \text{VaR}_p \mid X > \text{VaR}_p),$$

    where in both cases X represents a potential loss.
  - More sophisticated measures such as **expectiles** are also used.

Comments

- ☐ The T-year return level is often called 'the level exceeded once on average every T years', and is easily misinterpreted:
  - 'on average' does not mean that disasters arise at regular T-year intervals!
  - selection is often discounted — if M independent time series are monitored, then we expect M/T T-year events each year;
  - the assumption of stationarity is rarely true, so large events may cluster together in periods of elevated risk.
- ☐ Preferable to refer to quantiles — but probably impossible to change a cultural icon!
- ☐ Return levels and return periods are parameters of distributions, but future events are as-yet unobserved random variables, and it may be useful to consider their distributions. The distribution of the largest value X_T to be observed over T blocks of future background observations is $G^T(y)$, and it may be better to use this for risk analysis, in a Bayesian approach (later, probably).