## Statistics of Extremes

Anthony Davison

©2011

http://stat.epfl.ch

### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivation</td>
<td>2</td>
</tr>
<tr>
<td>Basic mathematical problem</td>
<td>3</td>
</tr>
<tr>
<td>Application areas.</td>
<td>15</td>
</tr>
<tr>
<td>Brief history</td>
<td>20</td>
</tr>
<tr>
<td>Why study extremes?</td>
<td>22</td>
</tr>
<tr>
<td>Resources</td>
<td>26</td>
</tr>
<tr>
<td>Course details</td>
<td>30</td>
</tr>
<tr>
<td>Maxima</td>
<td>31</td>
</tr>
<tr>
<td>Stability</td>
<td>32</td>
</tr>
<tr>
<td>Extremal types theorem</td>
<td>34</td>
</tr>
<tr>
<td>GEV</td>
<td>35</td>
</tr>
<tr>
<td>Summary</td>
<td>37</td>
</tr>
<tr>
<td>Extremal Types Theorem</td>
<td>38</td>
</tr>
<tr>
<td>Extremal types theorem</td>
<td>39</td>
</tr>
<tr>
<td>Tail quantile function</td>
<td>40</td>
</tr>
<tr>
<td>Helly–Bray theorem</td>
<td>41</td>
</tr>
<tr>
<td>Domains of attraction</td>
<td>46</td>
</tr>
<tr>
<td>von Mises conditions</td>
<td>47</td>
</tr>
<tr>
<td>Penultimate approximation</td>
<td>50</td>
</tr>
<tr>
<td>Limitations</td>
<td>51</td>
</tr>
<tr>
<td>Convergence</td>
<td>52</td>
</tr>
</tbody>
</table>

1
North Sea floods of 1953

- On the night of 31 January–1 February 1953, a combination of a high spring tide and a severe European windstorm caused a storm tide in the North Sea.
- The water level locally exceeded 5.6 metres above mean sea level, overwhelmed sea defences, and flooded areas of the Netherlands, England, Belgium, Denmark and France. Around 2500 people died.
- As a result, the Dutch government set up the so-called Delta Commission, which commissioned work that finished after nearly 50 years in 1997; another phase started in 1996 and will finish in 2015.
- The goal is to construct flood defences such that the acceptable period over which complete failure would occur is
  - North and South Holland (excluding wieingermeer): 1 per 10,000 years
  - Other areas at risk from sea flooding: 1 per 4,000 years
  - Transition areas between high land and low land: 1 per 2,000 years
- This is based on data observed over a much shorter period!

Statistics of extremes

Watersnoodramp, 1953

Statistics of extremes
Erith, London, 1953

Statistics of extremes

Autumn 2011 – slide 9

Thames Barrier

Statistics of extremes

Autumn 2011 – slide 10
Global Weirding?

Warning: extreme weather ahead
Tornadoes, wildfires, droughts and floods were once seen as freak conditions. But the environmental disasters now striking the world are shocking signs of "global weirding".

June Hulse, guardian.co.uk, Monday 13 June 2011 19:00 BST

Drought zones have been declared across much of England and Wales, yet Scotland has just registered its wettest-ever May. The wettest British spring in 100 years followed one of the coldest UK winters in 300 years. June in London has been colder than March. February was warm enough to whip on Snowdon, but last Saturday it snowed there.
Tanaguarena, 1999

- Following two weeks of intermittent rainfall, torrential rainfall on 14–16 December 1999 spawned landslides throughout the upper watersheds of the Cerro Grande River near the coast of Venezuela.
- Mud floods, debris flows and flood surges then destroyed much of Tanaguarena and other coastal tourist towns. Perhaps 30,000 people died.
- The data are from the airport at Maiquetia: the estimated recurrence time for the three-day rainfall is between 250 years and 6 million years!
- Similar events, fortunately with less loss of life, have occurred nearby.
- Again, we need to extrapolate beyond the existing data.

Rainfall at Maiquetia

Daily rainfall, 1961–1999 Venezuela

Statistics of extremes
Rainfall at Maiquetia

Daily rainfall, 1961–1999 Venezuela

Rainfall [mm]

Statistics of extremes

Tanaguarena

Statistics of extremes
Cerro Grande rivermouth

Comparison of Cerro Grande fan before and after the Dec. 1999 flood disaster.

Statistics of extremes Autumn 2011 – slide 19

Le ciel est tombé sur la tête des Tessinois

INTEMPÉRIES. Depuis mercredi, il est tombé plus de 300 litres d'eau par mètre carré sur le versant sud des Alpes. Dans certains endroits du Tessin, cela correspond au double des quantités enregistrées habituellement pour tout le mois de septembre. Le risque d'innondations n'est toutefois pas aisé. Malgré ces conditions, un seul blessé est à déplorer. Il s'agit d'un homme de 55 ans, un éclair est tombé près de lui alors qu'il se trouvait devant sa maison à côté d'une fontaine.

Bien que moins touchée, la Suisse romande a eu droit à son lot de précipitations. À Genève, il est tombé plus de 145 mm d'eau en l'espace de trois jours. Dans le canton de Vaud, 50 mm de plus pour la seule journée de samedi ont arrosé la région de la Dôle. Entre Saint-Péray et Villars-sous-Vens, le trafic a été perturbé par de l'eau sur la chaussée.

MALHEUREUX GOTHARD

Une voiture de boute a recouvert la ligne CFF, hier vers 12 h 40, entre Lavorgo et Plan excellent, interrompant la circulation tout l'après-midi. C’est la deuxième perturbation coup sur coup qui frappe le trafic au Gothard. La ligne avait été coupée déjà 102 coups durant 15 heures le lendemain, après la collision entre une maquinerie de charbon de 77 tonnes et quatre voilières remplies de ballast. Cet accident était survenu dans un tunnel entre Giornico et Lavorgo.

Malgré les routes coupées et le trafic perturbé, le Tessin a échappé au pire en fin de semaine.

Statistics of extremes Autumn 2011 – slide 20
The 2003 European heat wave was one of the hottest summers on record in Europe, especially in France. Temperatures exceeded 40°C in many places in northern Europe, and reached almost 50°C in Portugal.

It led to health crises in several countries and combined with drought to create a crop shortfall in Southern Europe.

It has been linked to forest fires in Portugal, to dried-up rivers, to buckling of train track in the UK, and to flash floods in the Alps.

Over 35,000 more people than usual are thought to have died in Europe that summer.

Some scenarios of climate change predict increases of up to 7°C in summer temperatures in northern Europe over the next century. How can we predict their likely consequences?
Swiss summer temperatures 2001–2005

The Basel Committee rules regulate the behaviour of banks and other financial institutions.

Part of this involves estimating a quantity known as the Value at Risk, which can be viewed mathematically as an extreme quantile of a distribution (though this is far from being the entire story!)

How should this be estimated from a series of daily percent returns, defined as

\[ Y_t = 100 \times \log(P_t/P_{t-1}) \]

where \( P_t \) is the closing price on day \( t \)?

How to measure the risk to a financial system of big changes in the values of several companies simultaneously?
Basic mathematical problem

- Simplest case: $X_1, \ldots, X_n \sim F$. Require accurate inferences on tail of $F$—must deal with extrapolation outside observed data.

- Key issues:
  - there are very few observations in the tail of the distribution;
  - estimates are often required beyond the largest observed data value;
  - standard data analysis/model-fitting techniques work well where the data have greatest density, but can be severely biased in estimating tail probabilities.

- Usual lack of physical or empirical basis for extrapolation leads to the **extreme value paradigm**: Base tail models on asymptotically-motivated distributions.

- Of course in practice we must ask: are the mathematical asymptotic arguments relevant to the real world?
Application areas

- **Environment**: sea-levels, river levels, rainfall, avalanches, temperatures, pollution
- **Finance**: financial time series; (re-)insurance
- **Internet traffic** and other communications networks
- **Material science**: strength of materials and structures
- **Reliability and survival analysis**: failure times for components and systems, and for people
- **Athletics**: record times
- **Microarrays**: significance levels for many similar tests
- **Seismology**: large earthquakes/tsunamis
- ... any setting where rare events dominate risks to people, structures, institutions, ...

Statistics of extremes

Brief history

- 1928: Foundations of asymptotic argument developed by Fisher and Tippett
- 1940s: Asymptotic theory unified and extended by Gnedenko and von Mises
- 1950s: Use of asymptotic distributions for statistical modelling by Gumbel and Jenkinson
- 1970s: Classic limit laws generalized by Pickands (and others), use of threshold methods in hydrology
- 1980s: Extension of theory to stationary processes, connections to theory of point processes and regular variation, use of regression techniques in extremal analysis
- 1990s: Development of multivariate models and other techniques as a means to improve inference; increasing applications to finance
- 2000s: Interest in space-time applications (climate change)
Founders
Ronald Alymer Fisher (1890–1962)
Leonard Henry Caleb Tippett (1902–1985)

Why study extremes?
- Many important applications!
- Rich probabilistic basis not often seen in other courses (regular variation, point processes, functional analysis, ...)
- Important (and in many cases unsolved) statistical problems, an area of research currently in very active development
- Fun for all!
Resources

Books
- Reiss and Thomas (2007) Statistical Analysis of Extreme Values, Birkhauser

R packages evd, evdbayes, evir, extRemes, fExtremes, POT, SpatialExtremes

Journal Extremes (published by Springer)

Projected course outline

Univariate data: maxima/minima (weeks 1–4)
Univariate data: point process approach (weeks 5–6)
Beyond the IID case: non-stationarity and dependence (weeks 7–8)
Multivariate extremes (weeks 9–12)
More complex settings (weeks 13–14)

Course details

Place: MA30 (= MA A3 30)
Lectures 8.15–10.00, Friday 23 September 2009 onwards
Exercises 13.15–15.00, Friday 23 September 2009 onwards
I shall also use material from elsewhere
Form of exam: written (60%) but with project component (40%)
Course material can be downloaded from

http://stat.epfl.ch/page-70154.html
Stability

Fundamental to all characterizations of extreme value processes and the basis for extrapolation is the concept of **stability**.

□ For example, we might propose one model for the annual maximum of a process, and another for the 5-year maximum. Since the 5-year maximum will be the maximum of 5 annual maxima, the models should be mutually consistent.

□ Similarly, a model for exceedances over a high threshold should remain valid (in a precise sense) for exceedances of higher threshold.

□ The expression of such stability requirements as mathematical statements leads to asymptotic models.

Probability framework for maxima

□ Let $X_1, \ldots, X_m \overset{\text{iid}}{\sim} F$ and define the maximum

$$M_m = \max\{X_1, \ldots, X_m\}.$$ 

Then the distribution function of $M_m$ is

$$\Pr\{M_m \leq x\} = \Pr\{X_1 \leq x, \ldots, X_m \leq x\} = \Pr\{X_1 \leq x\} \times \cdots \times \Pr\{X_m \leq x\} = F(x)^m.$$ 

□ $F$ is unknown, so approximate $F_m$ by some limit distribution.

□ What distributions can arise? As $m \to \infty$, obviously

$$F(x)^m \to \begin{cases} 0, & F(x) < 1, \\ 1, & F(x) = 1, \end{cases}$$

so $M_m \overset{D}{\to} x_F$, where $x_F = \sup\{x : F(x) < 1\}$ is the upper support point of $F$. The limit distribution is **degenerate**.
Reminder/Minima

**Definition 1** Let $X, X_1, X_2, \ldots$ be random variables with distribution functions $F, F_1, F_2, \ldots$. Then as $m \to \infty$, $X_m$ converges to $X$ in distribution (or in law), $X_m \overset{D}{\to} X$, if

$$\lim_{m \to \infty} F_m(x) = F(x) \quad \text{at every } x \text{ where } F(x) \text{ is continuous.}$$

This is the appropriate notion of convergence here, because we shall want to compare probabilities of events, and this gives us a way of doing so.

All the ideas apply equally to minima, because

$$\min(X_1, \ldots, X_m) = -\max(-X_1, \ldots, -X_m).$$

Our general discussion is for maxima, and we make this transformation without comment when we model minima.

Classical limit laws

We have seen that as $n \to \infty$, $M_n$ will converge in distribution to the degenerate law putting unit mass at $x_F$.

- But recall the Central Limit Theorem: under regularity conditions and with $\mu_m = \mu \in \mathbb{R}$ and $\sigma_m = \sigma / \sqrt{m} > 0$, then as $m \to \infty$,

$$\frac{X_m - \mu_m}{\sigma_m} \overset{D}{\to} N(0, 1):$$

rescaling is needed to obtain a non-degenerate limit.

- Same applies here: we seek limits of the rescaled quantity

$$\frac{M_m - b_m}{a_m}$$

for suitable sequences $\{a_m\} > 0$ and $\{b_m\} \subset \mathbb{R}$.

Examples

**Example 2** Find suitable sequences such that maxima of independent variables from the (a) exponential, (b) Frechét, and (c) uniform distributions have non-degenerate limiting distributions.
Example: Exponential maxima

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

Statistics of extremes

Example: Normal maxima

Maxima of standard normal variables also converge, with

\[
a_n = (2 \log n)^{-0.5}, \quad b_n = (2 \log n)^{0.5} - 0.5(2 \log n)^{-0.5}(\log \log n + \log 4\pi),
\]

but convergence is extremely slow.

Distributions of maxima and renormalized maxima of \( n = 1, 7, 30, 365, 3650 \) standard normal variables (from left to right), with Gumbel distribution (heavy).

Statistics of extremes

Max-stability

**Definition 3** A distribution \( H \) is said to be **max-stable** if

\[
H^k(x) = H(a_k x + b_k), \quad k = 1, 2, \ldots,
\]

for some constants \( a_k \) and \( b_k \).

A continuity argument can be used to show that if \( H \) is max-stable, then there exist functions \( a_t \) and \( b_t \) such that \( H^t(x) = H(b_t + a_t x) \) for all \( t \in \mathbb{R}_+ \).

**Lemma 4** If a limiting distribution \( H \) for rescaled maxima exists, it must be max-stable.
**Definition 5** The distributions $F$ and $F^*$ are of the **same type** if there are constants $a > 0$ and $b$ such that $F^*(ax + b) = F(x)$ for all $x$.

**Theorem 6 (Extremal types theorem)** If there exist sequences of constants $a_m > 0$ and $b_m$ such that, as $n \to \infty$,

$$\Pr\{(M_m - b_m)/a_m \leq x\} \to H(x)$$

for some non–degenerate distribution $H$, then $H$ has the same type as one of the following distributions:

- **I**: $H(x) = \exp\{-\exp(-x)\}$, $-\infty < x < \infty$;
- **II**: $H(x) = \begin{cases} 0, & x \leq 0, \\ \exp(-x^{-\alpha}), & x > 0, \alpha > 0; \end{cases}$
- **III**: $H(x) = \begin{cases} \exp\{-(x)^{\alpha}\}, & x < 0, \alpha > 0, \\ 1, & x \geq 0. \end{cases}$

Conversely, each of these $H$’s may appear as a limit for the distribution of $(M_m - b_m)/a_m$, and does so when $H$ itself is the distribution of $X$.

**Three limiting distributions**

- The three types are known as the **Gumbel**, **Fréchet** and **Weibull** (strictly, negative Weibull), respectively.
- The Fréchet (Type II) is bounded below, and the negative Weibull (Type III) is bounded above.
- The standard Weibull is a distribution for minima.
Using the limit law

- We assume that for some $a > 0$ and $b$,
\[
\Pr\{(M_m - b)/a \leq x\} \approx H(x),
\]
or equivalently,
\[
\Pr\{M_m \leq x\} \approx H(b + ax) = H^*(x),
\]
where $H^*$ is of the same type as $H$.

- That is, the family of extreme value distributions may be fitted directly to a series of observations of $M_m$.

- However, it is inconvenient to have to work with three possible limiting families.

Generalized extreme value distribution

- This family encompasses all three of the previous extreme value limit families:
\[
H(x) = \exp\left\{ - \left[ 1 + \xi \left( \frac{x - \eta}{\tau} \right) \right]^{-1/\xi} \right\},
\]
defined on $\{x : 1 + \xi(x - \eta)/\tau > 0\}$.

- From now on let $x_+ = \max(x, 0)$.

- $\eta$ and $\tau$ are location and scale parameters

- $\xi$ is a shape parameter determining the rate of tail decay, with
  - $\xi > 0$ giving the heavy-tailed (Fréchet) case
  - $\xi = 0$ giving the light-tailed (Gumbel) case
  - $\xi < 0$ giving the short-tailed (negative Weibull) case

Summary

- Statistics of extremes is an important domain of stochastics at the intersection of probability/stochastic modelling/statistics.

- A key element is extrapolation outside the range of the available data.

- Natural consistency arguments lead to the idea that limiting distributions, if they exist, must be max-stable.

- The extremal types theorem, analogous to the central limit theorem for sums, shows that the only max-stable distribution is the generalized extreme-value distribution, which is the basis of modelling for sample maxima and minima.
Extremal Types Theorem

Theorem 7 (Extremal types theorem) If there exist sequences of constants \(a_m > 0\) and \(b_m\) such that as \(m \to \infty\),

\[
\Pr\{(M_m - b_m)/a_m \leq x\} \to H(x)
\]

for some non-degenerate distribution \(H\), then

\[
H(x) = \exp\left\{-\left[1 + \xi \left(\frac{x - \eta}{\tau}\right)\right]^{-1/\xi}\right\},
\]

for some \(\eta, \xi \in \mathbb{R}\) and some \(\tau \in \mathbb{R}_+\), where we write \(x_+ = \max(x, 0)\), and with the case \(\xi = 0\) interpreted as the limit when \(\xi \to 0\).

Conversely, each of these \(H\)’s may appear as a limit for the distribution of \((M_m - b_m)/a_m\), and does so when \(H\) itself is the distribution of \(X\).

Tail quantile function

Definition 8 If \(F\) is a continuous distribution function with inverse function

\[
F^{-1}(u) = \inf\{x : F(x) \geq u\},
\]

then the corresponding tail quantile function is

\[
Q(y) = F^{-1}(1 - 1/y), \quad y \in [1, \infty].
\]

- Suppose independent observations from \(F\) arrive each day. Then \(F\{Q(y)\} = 1 - 1/y\) means that the level \(Q(y)\) will be exceeded over average once every \(y\) days.
- Note that \(Q(y)\) is monotone increasing, with \(Q(1) = \inf\{x : F(x) \geq 0\}, \quad Q(\infty) = \inf\{x : F(x) \geq 1\}\) being the lower and upper terminals of \(F\).
Helly–Bray theorem

**Theorem 9 (Helly–Bray)** Let $X, X_1, X_2, \ldots$ be random variables with distribution functions $F, F_1, F_2, \ldots$. Then $X_n \xrightarrow{D} X$ as $n \to \infty$ if and only if for every real bounded continuous function $g$

$$E\{g(X_n)\} \to E\{g(X)\} \text{ or equivalently } \int g(x) \, dF_n(x) \to \int g(x) \, dF(x).$$

Thus the limit random variable $X$ is constant iff $E\{g(X_n)\} \to g(c)$ for all such $g$.

We will use this result to show that when renormalized, $M_m = \max\{X_1, \ldots, X_m\}$ converges to a limiting distribution $H$, by showing that for every suitable $g$,

$$E \left\{ g \left( \frac{M_m - b_m}{a_m} \right) \right\} \to \int g(y) \, dH(y),$$

where $H$ is of extreme-value form.

We will impose the following condition on $Q$:

$$(C) : \exists a : \mathbb{R}_+ \to \mathbb{R}_+ \text{ such that } \lim_{x \to \infty} \frac{Q(ux) - Q(x)}{a(x)} = q(u), \quad u > 0$$

exists.

### Possible limits under $(C)$

**Lemma 10** The only possible limits in $(C)$ are

$$q_\xi(u) = c \frac{u^\xi - 1}{\xi}, \quad c \geq 0, \xi \in \mathbb{R},$$

where we take $q_0(u) = c \log u$.

### Domains of attraction

- What limit (if any) will arise for maxima from a given distribution function $F$?
- The **domain of attraction** of the Gumbel distribution $\exp\{-\exp(-y)\}$ is the class of all distributions $F$ such that if $X_1, \ldots, X_m \sim F$ and $M_m = \max(X_1, \ldots, X_m)$, then sequences $\{b_m\}$ and $\{a_m\}$ exist such that the limiting distribution of $(M_m - b_m)/a_m$ is Gumbel.
- Domains of attraction for other GEV distributions are defined similarly.
- Books such as Embrechts et al. (1997) and Beirlant et al. (2004) give discussions of necessary and sufficient conditions for $F$ to lie in different domains of attraction.
- From a statistical viewpoint this is not so useful: in practice we are never sure of $F$ and so will also be uncertain about the limiting distribution.
von Mises conditions

□ How do we determine sequences $a_m$ and $b_m$ and the limit distribution $H$ (if any)?
□ Simple sufficient (but not necessary) conditions are the von Mises conditions.
□ For a sufficiently smooth distribution $F'$ with upper terminal $x_F$, define the reciprocal hazard function as

$$r(x) = \frac{1 - F(x)}{f(x)}.$$ 

Then with

$$b_m = F^{-1}(1 - 1/m), \quad a_m = r(b_m), \quad \xi = \lim_{x \to x_F} r'(x)$$

the limit distribution of $(M_m - b_m)/a_m$ is GEV with shape parameter $\xi$.

□ The value $b_m = F^{-1}(1 - 1/m)$ is sometimes called the largest typical value of $F$; note that if its expectation exists, the largest order statistic satisfies

$$E(X_{(m)}) = F^{-1}\{m/(m + 1)\} \approx F^{-1}\{1 - 1/(m + 1)\} \approx b_m,$$

so $b_m$ is approximately the mean of the maximum of $X_1, \ldots, X_m$.

Example 11 Use the von Mises conditions to check the limiting distributions of maxima of the uniform, exponential, Fréchet, and normal distributions.

Penultimate approximation

□ Taking $\xi_m = r'(b_m)$ may give a better approximation to the distribution of $(M_m - b_m)/a_m$ for finite $m$ than does using the limiting approximation.
□ For the normal distribution, for example

$$\xi_7 = -0.324, \quad \xi_{30} = -0.176, \quad \xi_{365} = -0.097, \quad \xi_{3650} = -0.068, \quad \xi_{36500} = -0.052,$$

so the distribution of $M_m$ is short-tailed ($\xi < 0$) relative to the Gumbel limit even for very large $m$.

□ The statistical consequence is that even when we are quite sure that we should obtain a Gumbel limit, $\xi = 0$, we may prefer to fit the GEV with arbitrary $\xi$, in order to get a better approximation for finite $m$.

□ The graph on the following page shows the short tails very clearly as downward curvature, better modelled by taking $\xi < 0$ than by taking $\xi = 0$.

□ Transformation of the maxima (e.g. taking logs) may improve the closeness of the GEV approximation, at the cost of giving greater uncertainty when making predictions.
Normal example

100 replicates of renormalized normal maxima with $m = 7, 30, 365, 3650$

---

Limitations of the GEV limit

The GEV limit law requires careful reading.

☐ It states that if linearly renormalized maxima have a limiting distribution, then that limit will be a member of the GEV family. It does not guarantee the existence of a limit, and there are many classes of models for which no limit exists.

☐ One is the case $F(x) = 1 - 1/\log x$, $x \geq e$, for which the tail is too heavy for linear renormalization to work.

☐ Another is the Poisson distribution. If $\{X_i\}$ is a sequence of Poisson variables with mean $\lambda$, a sequence of integer constants $\{I_m\}$ can be found such that

$$\lim_{m \to \infty} \Pr \{M_m = I_m \text{ or } I_m + 1\} = 1$$

which implies that no limit distribution for renormalized maxima exists.

☐ The discreteness in the Poisson distribution causes the oscillation between $I_m$ and $I_m + 1$. It is not the cause of the limit degeneracy, which is actually a consequence of the Poisson tail decay.

---

Statistics of extremes Autumn 2011 – slide 60

---

Statistics of extremes Autumn 2011 – slide 61
Convergence and approximation

- There has been a good deal of work on the speed of convergence of $M_m$ to the limiting regime, which depends on the underlying distribution $F$—we have seen, for example, that convergence is slow for maxima of $m$ Gaussian variables.

- From a statistical viewpoint, this is not so useful: we use the GEV as an approximate distribution for sample maxima for finite (small?) $m$, so the key question is whether the GEV fits the available data—assess this empirically.

- Direct use of the GEV rather than the three types separately allows for flexible modelling, and ducks the question of which type is most appropriate—the data decide.

- Testing for fit of one type or another is usually unhelpful, because setting (say) $\xi = 0$ can give unrealistically precise inferences. Often uncertainty in extremes is (appropriately) large, and it can be misleading to constrain inferences artificially.

Summary

- Have seen a ‘proof’ of the extremal types theorem.
- Have given simple sufficient conditions for convergence of maxima.
- Have discussed penultimate approximations and their consequences.
- Have seen some examples where the theorem doesn’t hold.
- Next time: statistical use of the extremal types theorem.