The Modeling of Rare Events: from Methodology to Practice and Back

Paul Embrechts

Department of Mathematics

Director of RiskLab, ETH Zurich

Senior SFI Chair

www.math.ethz.ch/~embrechts

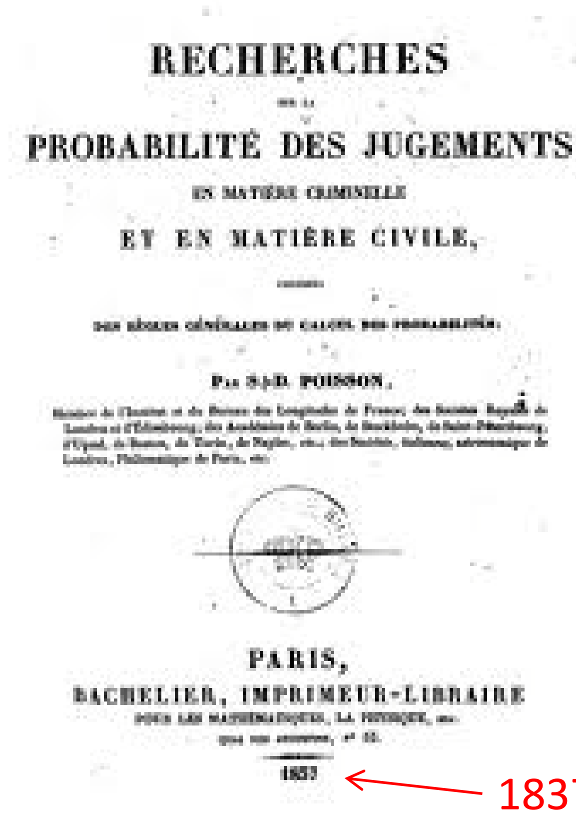
Summary:

- A bit of history
- A bit of theory
- An application
- Further work

Perhaps the first:

Nicolaus Bernoulli (1687 – 1759) who, in 1709, considered the actuarial problem of calculating the mean duration of life of the last survivor among n men of equal age who all die within t years. He reduced this question to the following:
n points lie at random on a straight line of length t, calculate the mean largest distance from the origin.

Often quoted as the start:



RECHERCHES

PROBABILITÉ DES JUGEMENTS

EN MATIÈRE CRIMINELLE

ET EN MATIÈRE CIVILE,

DES RÈGLES GÉNÉRALES DU CALCUL DES PROBABILITÉS.

Par S.-D. POISSON,

PARIS,

BACHELIER, IMPRIMEUR-LIBRAIRE

1837

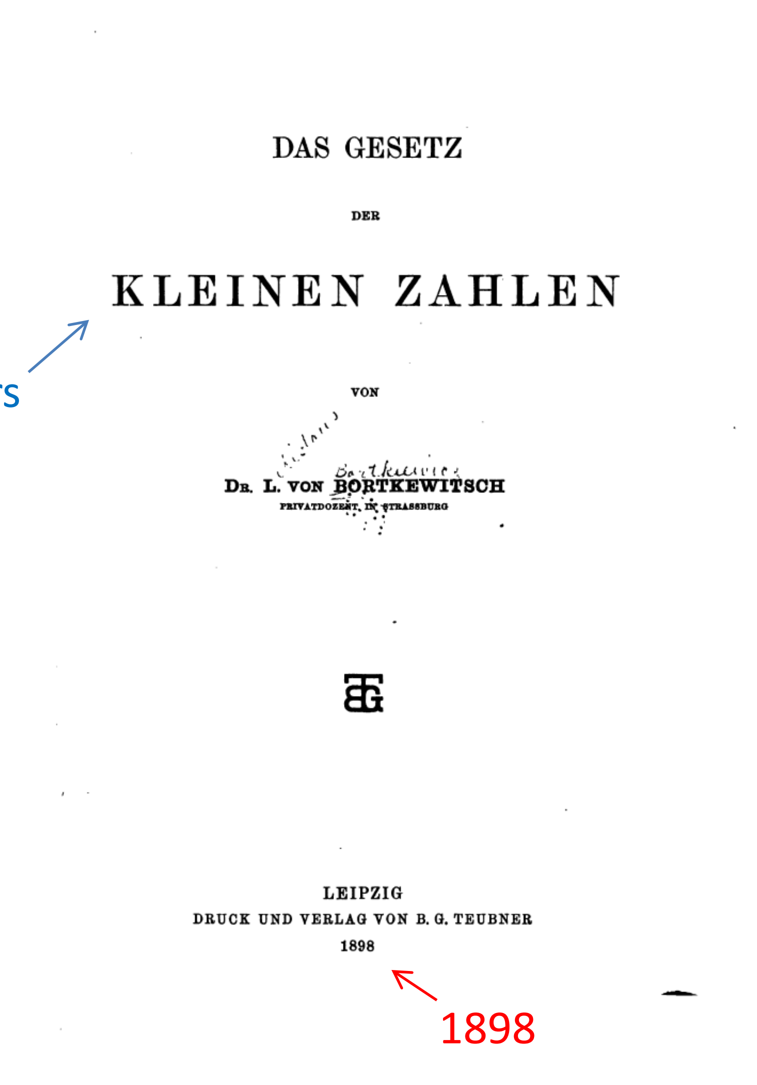
1837 (p 206)

Simon Denis Poisson (1781 – 1840)

(however Cotes (1714), de Moivre (1718), ...)

However, the real start with relevance to EVT was given by:

The Law of Small Numbers



DAS GESETZ

DER

KLEINEN ZAHLEN

VON

DR. L. VON BORTKEWITSCH

PRIVATDOZENT IN STRASSBURG

LEIPZIG

DRUCK UND VERLAG VON B. G. TEUBNER

1898

1898

Ladislaus J. von Bortkiewicz

(1868 – 1931)

(Prussian army horse-kick data)

Developments in the early to mid 20th century:

R.A. Fisher

L.H.C. Tippett

M.R. Fréchet

E.H.W. Weibull

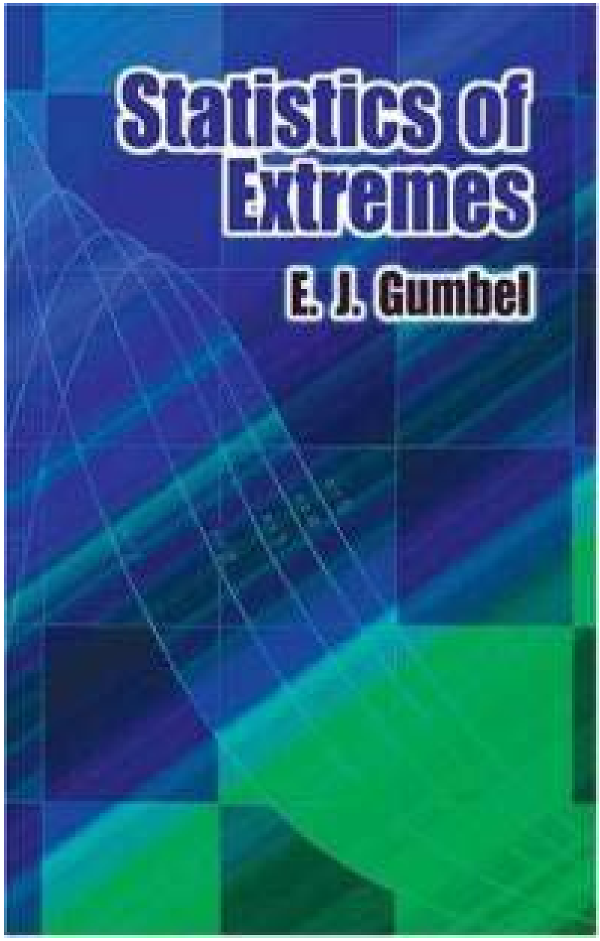
E.J. Gumbel

R. von Mises

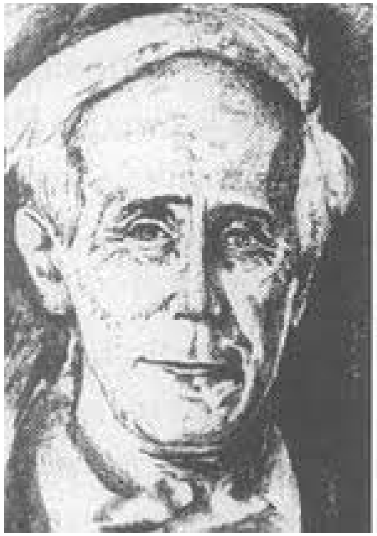
B.V. Gnedenko

· · ·

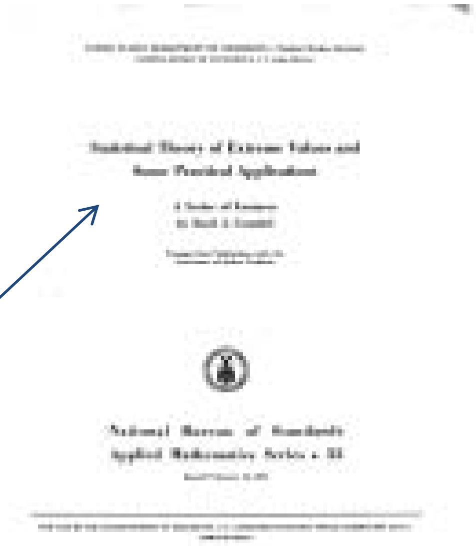
With an early textbook summary:



(1958)



Emil Julius Gumbel
(1891 – 1966)

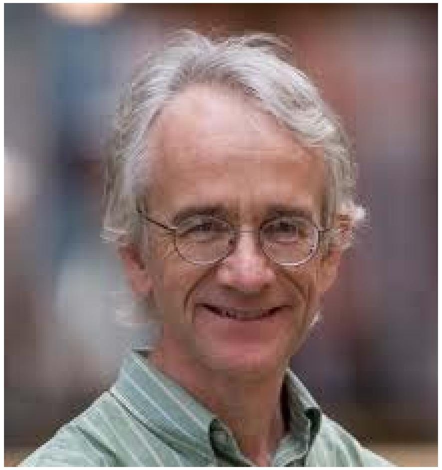


Statistical Theory of Extreme Values and Some Practical Applications.
National Bureau of Standards, 1954

Then the later-20th Century explosion:

Laurens de Haan

Sidney I. Resnick



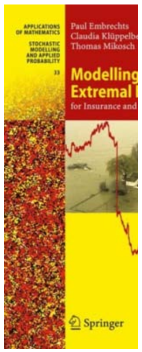
Richard L. Smith

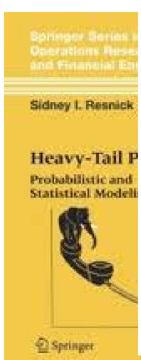
M. Ross Leadbetter

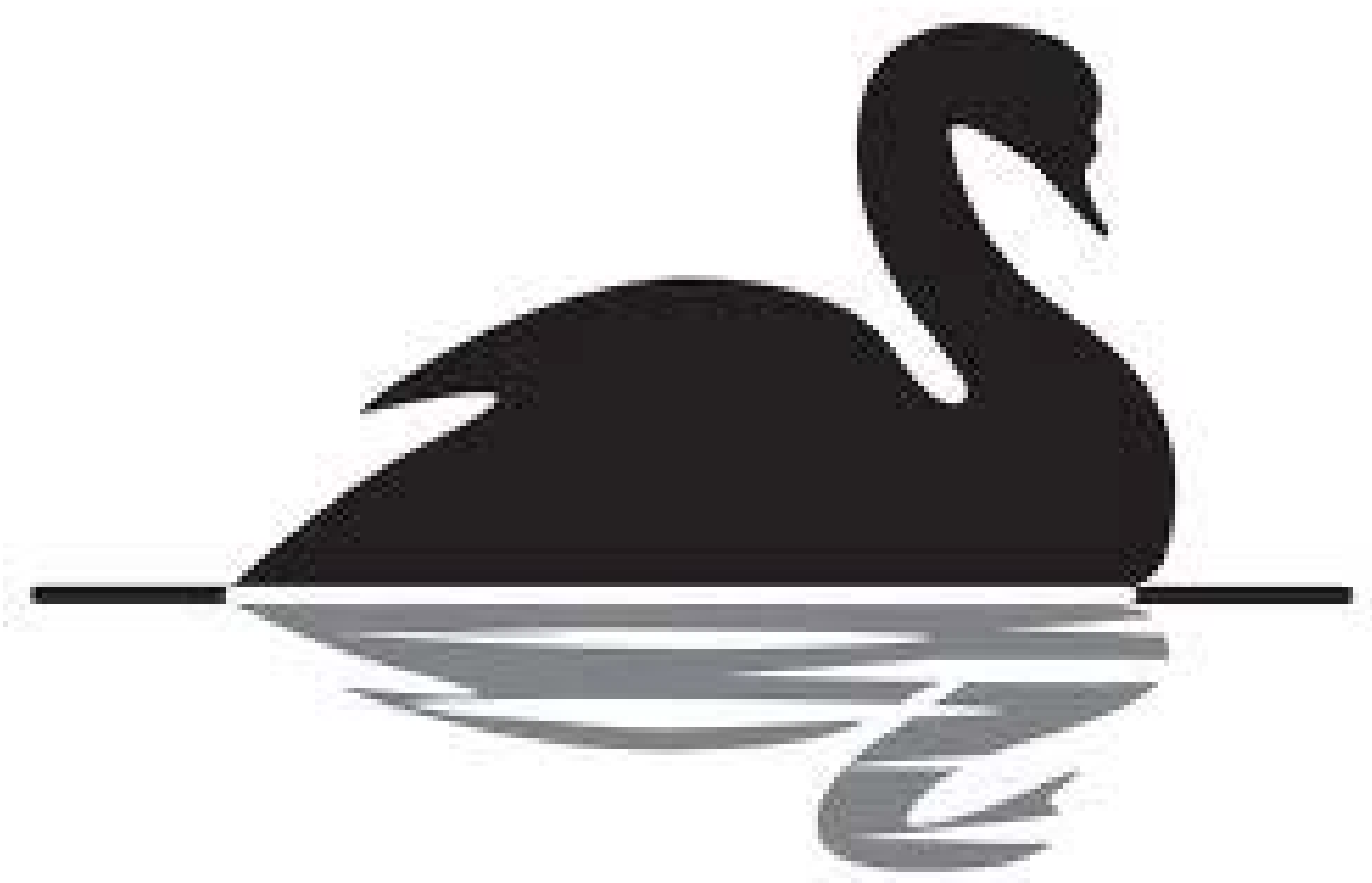
and so many more ...

Black Swan:

and Reflected inary







o o o

Recall the Central Limit Theorem:

Limiting Behaviour of Sums or Averages

(See [Embrechts et al., 1997], Chapter 2.)

We are familiar with the central limit theorem.

Let $X_1, X_2, \ldots$ be iid with finite mean μ and finite variance σ^2. Let $S_n = X_1 + X_2 + \ldots + X_n$. Then

$$P\left((S_n - n\mu) / \sqrt{n\sigma^2} \leq x \right) \stackrel{n\to\infty}{\longrightarrow} \Phi(x) \, ,$$

where Φ is the distribution function of the standard normal distribution

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du \, .$$

Note, more generally, the limiting distributions for appropriately normalized sample sums are the class of α–stable distributions; Gaussian distribution is a special case.

The basic 1-d EVT set-up, the «Extreme Value Theorem»:

EVT = Extreme Value Theory
1-d = one dimensional

Limiting Behaviour of Sample Extrema

(See [Embrechts et al., 1997], Chapter 3.)

Let $X_1, X_2, \ldots$ be iid from F and let $M_n = \max(X_1, \ldots, X_n)$.

Suppose we can find sequences of real numbers $a_n > 0$ and b_n such that $(M_n - b_n)/a_n$, the sequence of normalized maxima, converges in distribution, i.e.

$$P\left((M_n - b_n)/a_n \leq x\right) = F^n(a_n x + b_n) \overset{n \to \infty}{\longrightarrow} H(x),$$

for some non–degenerate df $H(x)$.

If this condition holds we say that F is in the *maximum domain of attraction* of H, abbreviated $F \in \text{MDA}(H)$. Note that such an H is determined up to location and scale (by the convergence-to-types theorem), i.e. H specifies a unique *type* of distribution.

F2. Generalized Extreme Value Distribution

The general form of a df in the generalized extreme value (GEV) family is

$$H_\xi(x) = \begin{cases} \exp\left(-(1+\xi x)^{-1/\xi}\right) & \xi \neq 0\,, \\ \exp\left(-e^{-x}\right) & \xi = 0, \end{cases}$$

where $1 + \xi x > 0$ and ξ is the *shape* parameter. Note, this parametrization is continuous in ξ. For

$\xi > 0$ H_ξ is equal in type to classical Fréchet df

$\xi = 0$ H_ξ is equal in type to classical Gumbel df

$\xi < 0$ H_ξ is equal in type to classical Weibull df.

We introduce *location and scale* parameters μ and $\sigma > 0$ and work with $H_{\xi,\mu,\sigma}(x) := H_\xi((x-\mu)/\sigma)$. Clearly $H_{\xi,\mu,\sigma}$ is of type H_ξ.

F3. Fisher–Tippett Theorem (1928)

Theorem: If $F \in \text{MDA}(H)$ then H is of the type H_ξ for some ξ.

"If suitably normalized maxima converge in distribution to a non–degenerate limit, then the limit distribution must be an extreme value distribution."

Remark 1: Essentially all commonly encountered *continuous* distributions are in the maximum domain of attraction of an extreme value distribution.

Remark 2: We can always choose normalizing sequences a_n and b_n so that the limit law H_ξ appears in standard form (without relocation or rescaling).

When does $F \in \text{MDA}(H_\xi)$ hold?

1. Fréchet Case: ($\xi > 0$)

Gnedenko (1943) showed that for $\xi > 0$,

Power law

Beware!

$$F \in \text{MDA}(H_\xi) \iff 1 - F(x) = x^{-1/\xi} L(x) \text{ for } x > 0,$$

for some function $L(x)$ which is *slowly varying at* ∞.

A positive function L on $(0, \infty)$ is slowly varying at ∞ if $\forall t > 0 \quad \lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1$. Example: $L(x) = \ln(x)$.

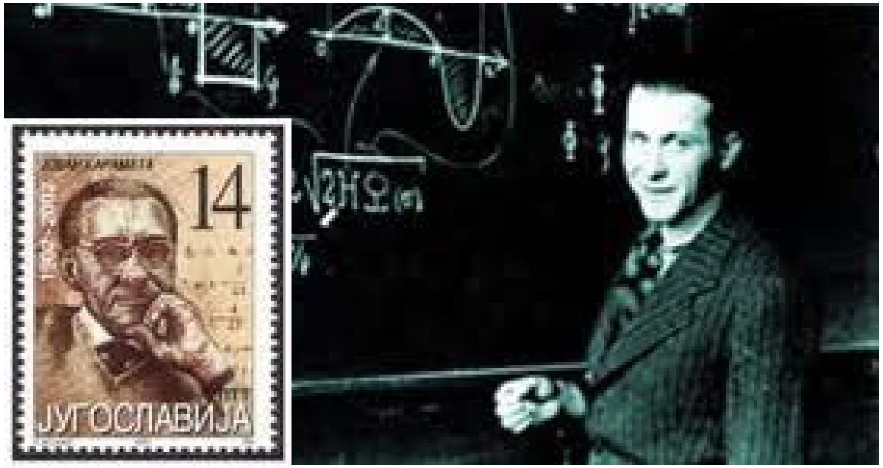
Summary:

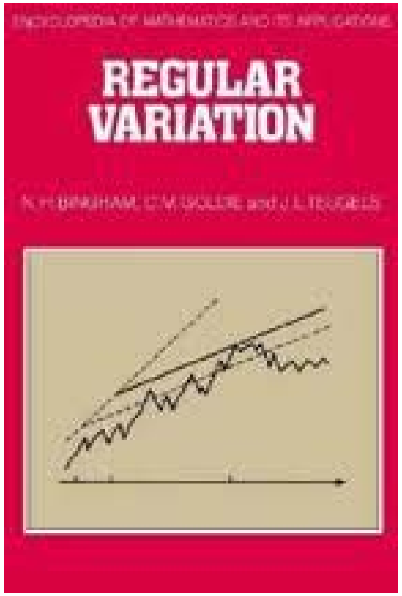
If the tail of the df F decays like a power function, then the distribution is in $\text{MDA}(H_\xi)$ for $\xi > 0$.

An interludium on Regular Variation, more in particular, on the Slowly Varying L in Gnedenko's Theorem:

One further name and a book:

Jovan Karamata
(1902 -1967)





by N.H. Bingham, C.M. Goldie and J.L. Teugels (1987): contains ALL about L-functions!

EVT and the POT method

Some isues:

Practice is too often frequency oriented ...

- every so often (rare event)
- return period, 1 in x-year event
- Value-at-Risk (VaR) in financial RM

... rather than more relevant severity orientation

- what if
- loss size given the occurence of a rare event
- Expected Shortfall E[X I X > VaR]

This is not just about theory but a RM attitude!

The Peaks Over Threshold (POT) Method

The excess distribution

Crucial point!

Given that a loss exceeds a *high threshold*, by how much can the threshold be exceeded?

Let u be the high threshold and define the *excess distribution* above the threshold u to have the df

$$F_u(x) = P(X - u \leq x \mid X > u) = \frac{F(x+u) - F(u)}{1 - F(u)},$$

for $0 \leq x < x_F - u$ where $x_F \leq \infty$ is the right endpoint of F.

Extreme value theory suggests the GPD is a *natural approximation* for this distribution.

Asymptotics of Excess Distribution

Theorem. (Pickands–Balkema–de Haan (1974/75)) We can find a positive, measurable function $\beta(u)$ such that

$$\lim_{u \to x_F} \sup_{0 \le x < x_F - u} \left| F_u(x) - G_{\xi,\beta(u)}(x) \right| = 0\,,$$

if and only if $F \in \text{MDA}\,(H_\xi)$, $\xi \in \mathbb{R}$.

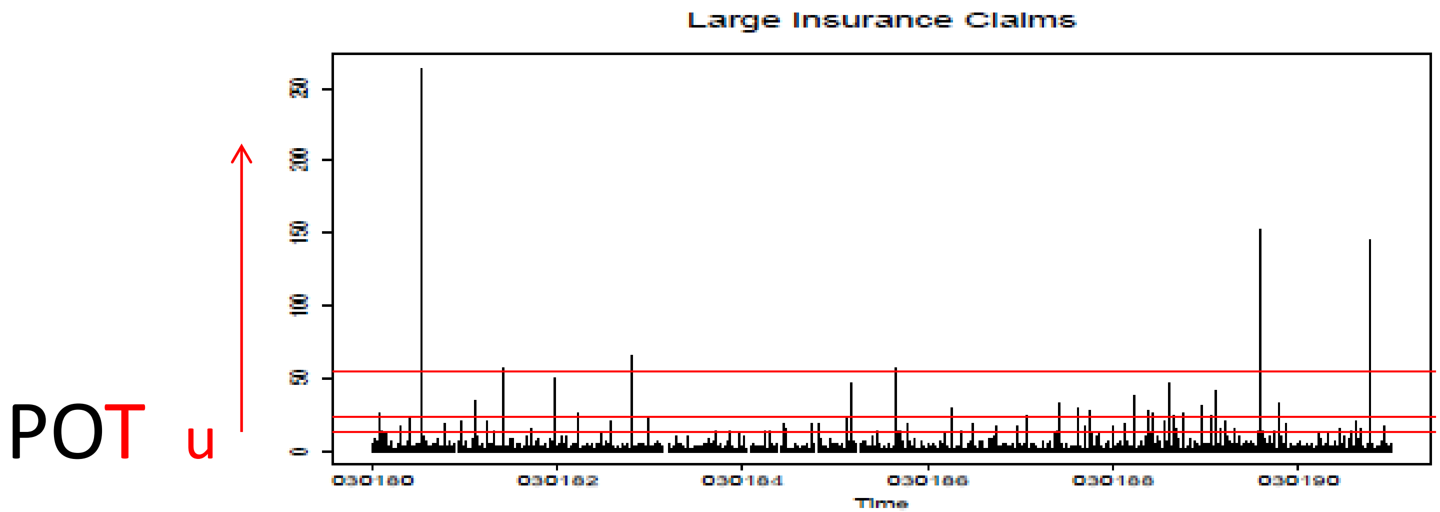
- The GPD is a two parameter distribution with df

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)_+^{-1/\xi} & \xi \neq 0\,, \\ 1 - \exp(-x/\beta) & \xi = 0\,, \end{cases}$$

where $\beta > 0$ and $a_+ = \max(a, 0)$, so the support is $x \ge 0$ when $\xi \ge 0$ and $0 \le x \le -\beta/\xi$ when $\xi < 0$.

Danish Fire Loss Example

The Danish data consist of 2167 losses exceeding one million Danish Krone from the years 1980 to 1990. The loss figure is a total loss for the event concerned and includes damage to buildings, damage to contents of buildings as well as loss of profits. The data have been adjusted for inflation to reflect 1985 values.



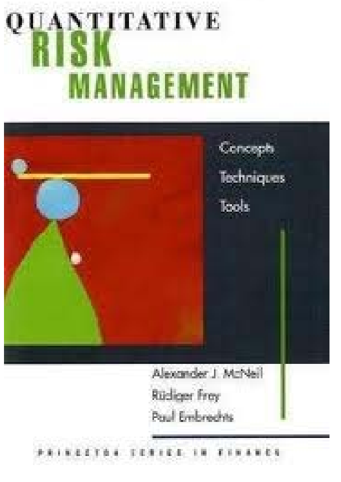
Start an EVT-POT analysis:

- First diagnostic checking
- Statistical techniques
- Graphical techniques
- Standard software in all relevant hard- and software environments: R, S-Plus, ...
- In our example,

  McNeil's QRM-LIB from:

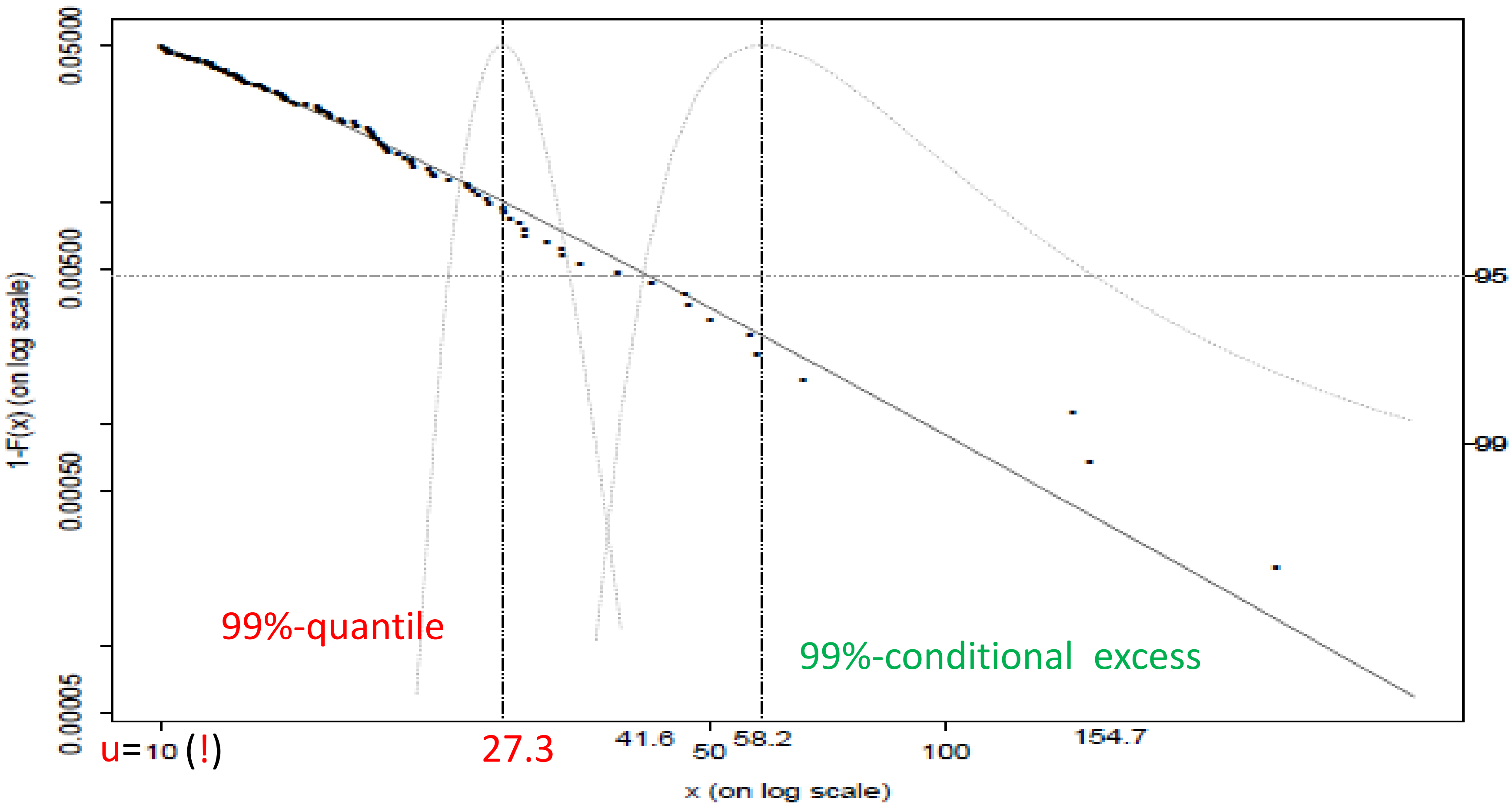
  A.J. McNeil, R.Frey and

  P. Embrechts (2005)



99%-quantile with 95% aCI (Profile Likelihood):
27.3 (23.3, 33.1)

99% Conditional Excess: E(X I X > 27.3) with aCI



A warning on slow convergence!

$$\lim_{u\uparrow x_0} \underbrace{\sup_{x\in(0,x_0-u)} \left|F_u(x) - G_{\xi,\beta(u)}(x)\right|}_{=:d(u)} = 0$$

Rate of convergence to the GPD for different distributions, as a function of the threshold u

L matters!

Distribution	Parameters	$\overline{F}$	$d(u)$
Exponential(λ)	$\lambda > 0$	$e^{-\lambda x}$	0
Pareto(α)	$\alpha > 0$	$x^{-\alpha}$	0
Double exp. parent		e^{-e^x}	$O(e^{-u})$
Student t	$\nu > 0$	$\overline{t}_\nu(x)$	$O(\frac{1}{u^2})$
Normal(0, 1)		$\overline{\Phi}(x)$	$O(\frac{1}{u^2})$
Weibull(τ, c)	$\tau \in \mathbb{R}_+ \setminus \{1\}, c > 0$	$e^{-(cx)^\tau}$	$O(\frac{1}{u^\tau})$
Lognormal(μ, σ)	$\mu \in \mathbb{R}, \sigma > 0$	$\overline{\Phi}(\frac{\log x - \mu}{\sigma})$	$O(\frac{1}{\log u})$
Loggamma(γ, α)	$\alpha > 0, \gamma \neq 1$	$\overline{\Gamma}_{\alpha,\gamma}(x)$	$O(\frac{1}{\log u})$
g-and-h	$g, h > 0$	$\overline{\Phi}(k^{-1}(x))$	$O(\frac{1}{\sqrt{\log u}})$

Some issues:

- Extremes for discrete data: special theory
- No unique/canonical theory for multivariate extremes because of lack of standard ordering, hence theory becomes context dependent
- Interesting links with rare event simulation, large deviations and importance sampling
- High dimensionality, $d > 3$ or 4 (sic)
- Time dependence (processes), non-stationarity
- Extremal dependence (← financial crisis)
- And finally ... APPLICATIONS ... COMMUNICATION !!

Thank you!