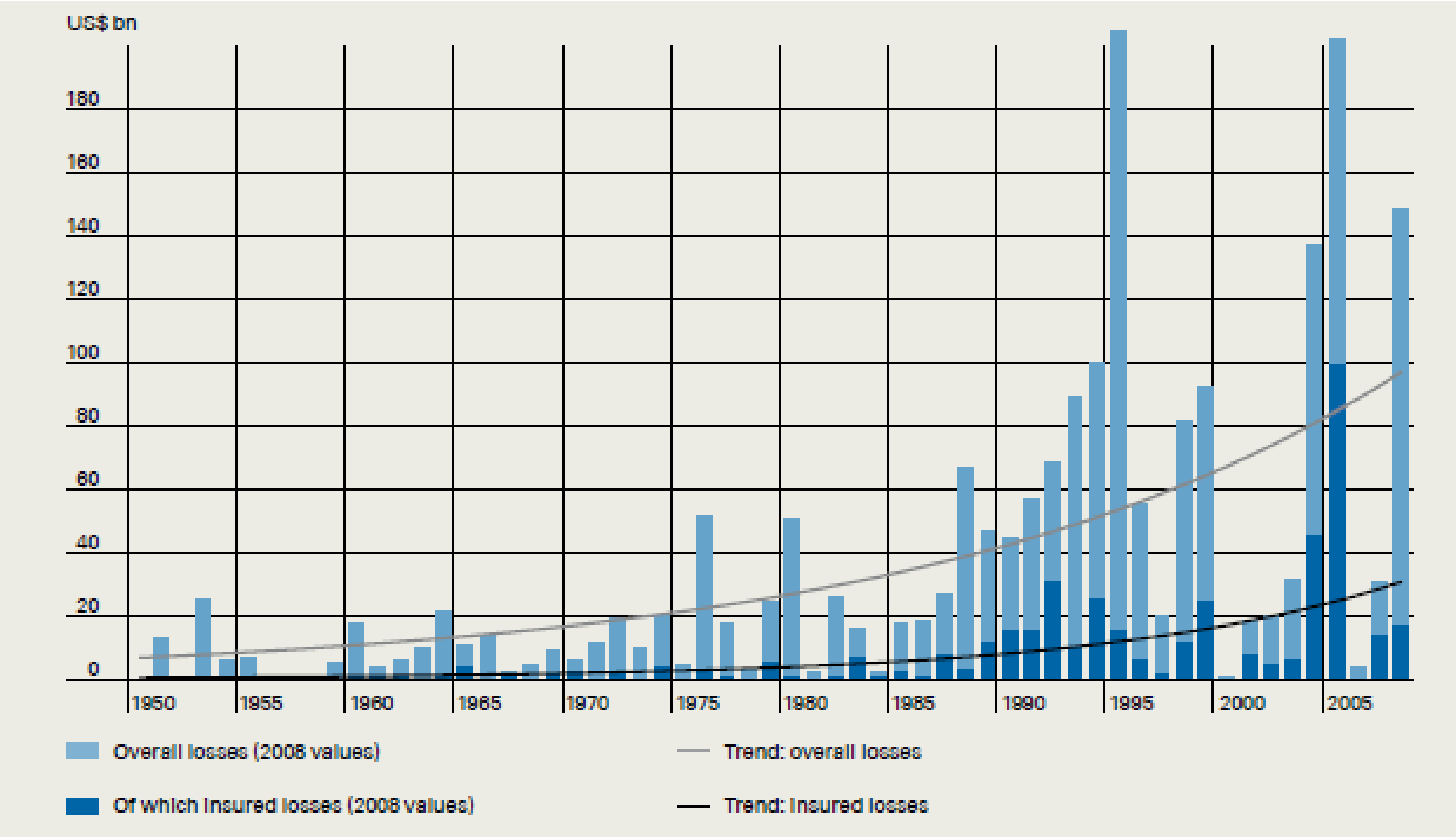
Application of Peaks Over Threshold method in insurance

Branka Skelin
Actuarial department
Agram life insurance

Croatian Quants Day
Department of Mathematic, University of Zagreb
Zagreb, May 6, 2011

- Peaks Over Threshold (POT) method
- Basic and generalized POT model
- Application of POT model

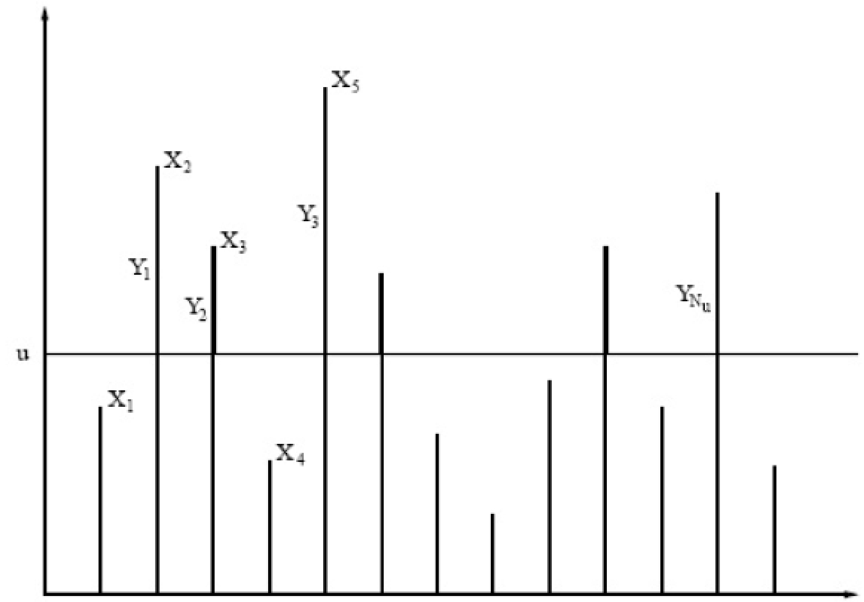
Peaks Over Threshold (POT) method



Overall losses and insured losses from great natural catastrophes, 1950.-2008.
Münich Re, Natural catastrophes 2008, Analyses, assessments, positions

Excesses over threshold *u*

$$F_u(x) = P\{X - u \le x \mid X > u\}$$



Generalized Pareto distribution

$$G_\xi(x) = \begin{cases} 1 - (1+\xi x)^{-1/\xi} & if \quad \xi \neq 0\,, \\ 1 - e^{-x} & if \quad \xi = 0\,, \end{cases}$$

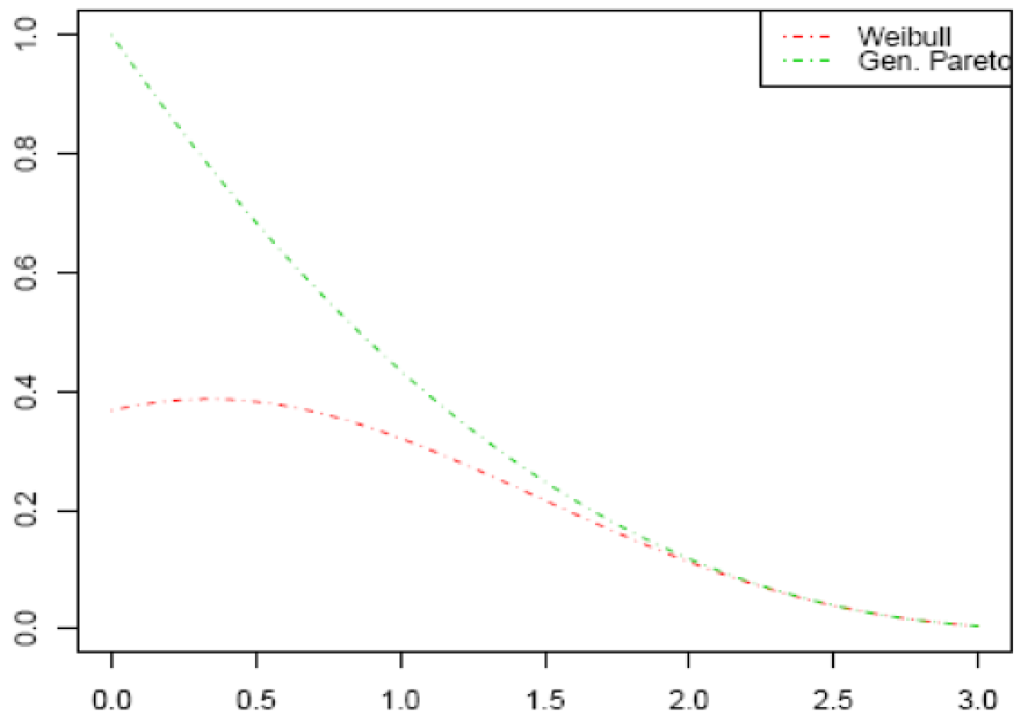
where

$$\begin{array}{ll} x \geq 0 & if \quad \xi \geq 0\,, \\ 0 \leq x \leq -1/\xi & if \quad \xi < 0\,. \end{array}$$

Balkema-de Haan-Pickands theorem

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

Weibull density and generalized Pareto density



Weibull density with parameter ξ=-0.3

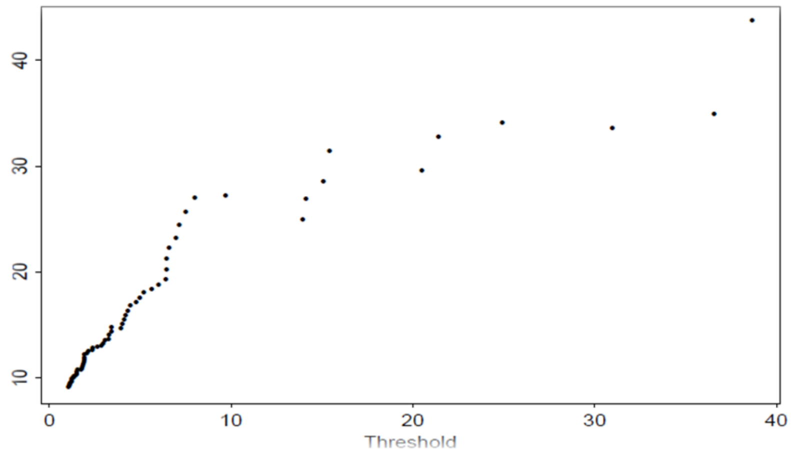
POT method

- **Data should be from one of the heavy tailed distribution**

  – QQ plot, sample mean excess plot

- **Threshold**

  – sample mean excess plot



Danish fire insurance losses from 1980. to 1990., losses over 1 mil. Danish Krone

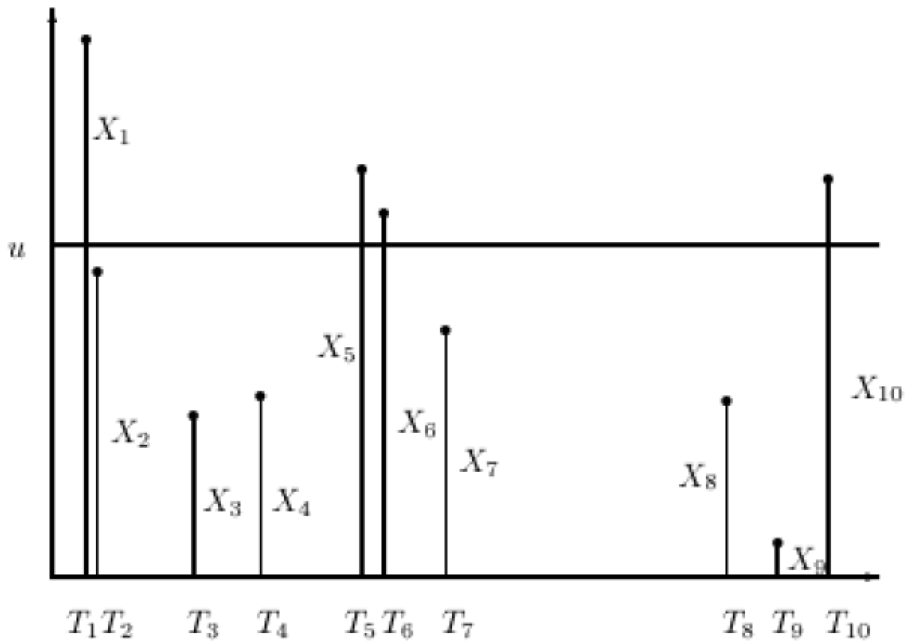
- **Parametar estimation**

  - maximum likelihood method – recommended for $\xi \geq 0.5$,

  - method of probability-weighted moments – recommended for small sample sizes and for $\xi \in (0, 0.4)$

Basic and generalized POT model

Observations above threshold



Occurrence times – modeled with Poisson process

Exceedances – modeled with generalized Pareto distribution

Basic POT model

Two dimensional point process $\{(T_i, X_i), 1 \leq i \leq N, X_i > u\}$ on $[-T, 0] \times (u, \infty)$

Exceedances

Distribution function of exceedances

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

↳ **generalized Pareto distribution**

⟹ (X_i - u) iid from generalized Pareto distribution

Crude residuals* can be defined as $W_i = \frac{1}{\xi} \log \left(1 + \xi \frac{X_i - u}{\sigma + \xi(u - \mu)} \right)$

⟹ W_i should be iid unit exponentially distributed

*Cox & Snell (1968)

Occurrence times

Homogeneous Poisson process with a constant intensity

$$\lambda = \left(1 + \xi \frac{u - \mu}{\sigma}\right)^{-1/\xi}$$

Rescaled inter-exceedance times can be defined with

$$Z_k = \lambda(T_k - T_{k-1}), \quad k = 1, ..., N + 1$$

where $T_0 = -T$ i $T_{N+1} = 0$

$\Longrightarrow$ Z_k should be unit exponentially distributed.

Generalized POT model

Model 1: $\mu(t) = \alpha_1 + \beta_1 t$ linear growth in μ

Model 2: $\sigma(t) = \exp(\alpha_2 + \beta_2 t)$ exponential growth in σ

Model 3 $\xi(t) = \alpha_3 + \beta_3 t$ linear growth in ξ

We can consider all possible combinations.

Rescaled inter-exceedence times

$$\lambda(t) = \left(1 + \xi(t)\frac{u - \mu(t)}{\sigma(t)}\right)^{-\frac{1}{\xi(t)}}$$

$$Z_k = \int_{T_{k-1}}^{T_k} \lambda(s)ds, \qquad k = 1, \ldots, N-1$$

Residuals $W_k = \frac{1}{\xi(T_k)} \ln\left(1 + \xi(T_k)\frac{X_k - u}{\sigma(T_k) + \xi(T_k)(u - \mu(T_k))}\right)$

Z_k and W_k should be unit exponentially distributed.

Loss number N_k above threshold u has a Poisson distribution with mean

$$\int_{365(k-1)}^{365k} \lambda(t)dt$$

N_k new losses in the year k are $Y_{k,1}, \ldots, Y_{k,N_k}$

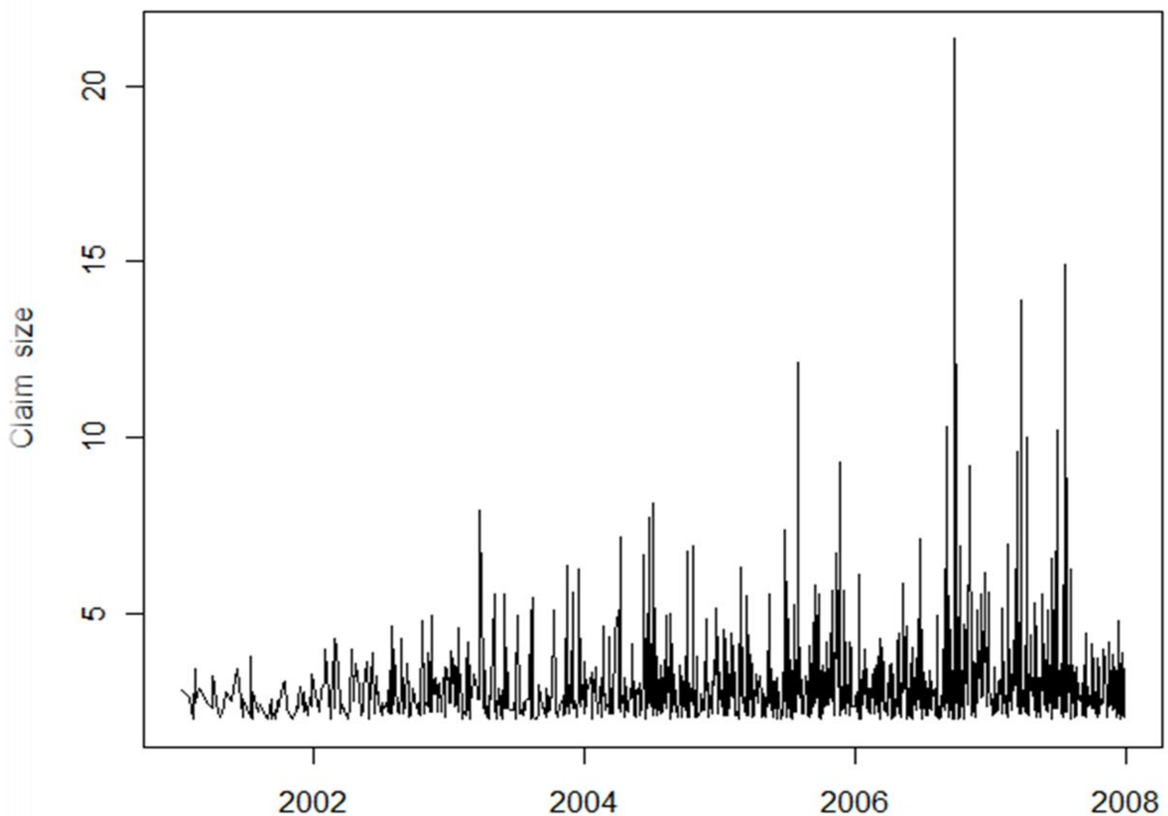
Excesses $(Y_{k,i} - u)$ have a GPD

Total loss volume in year k

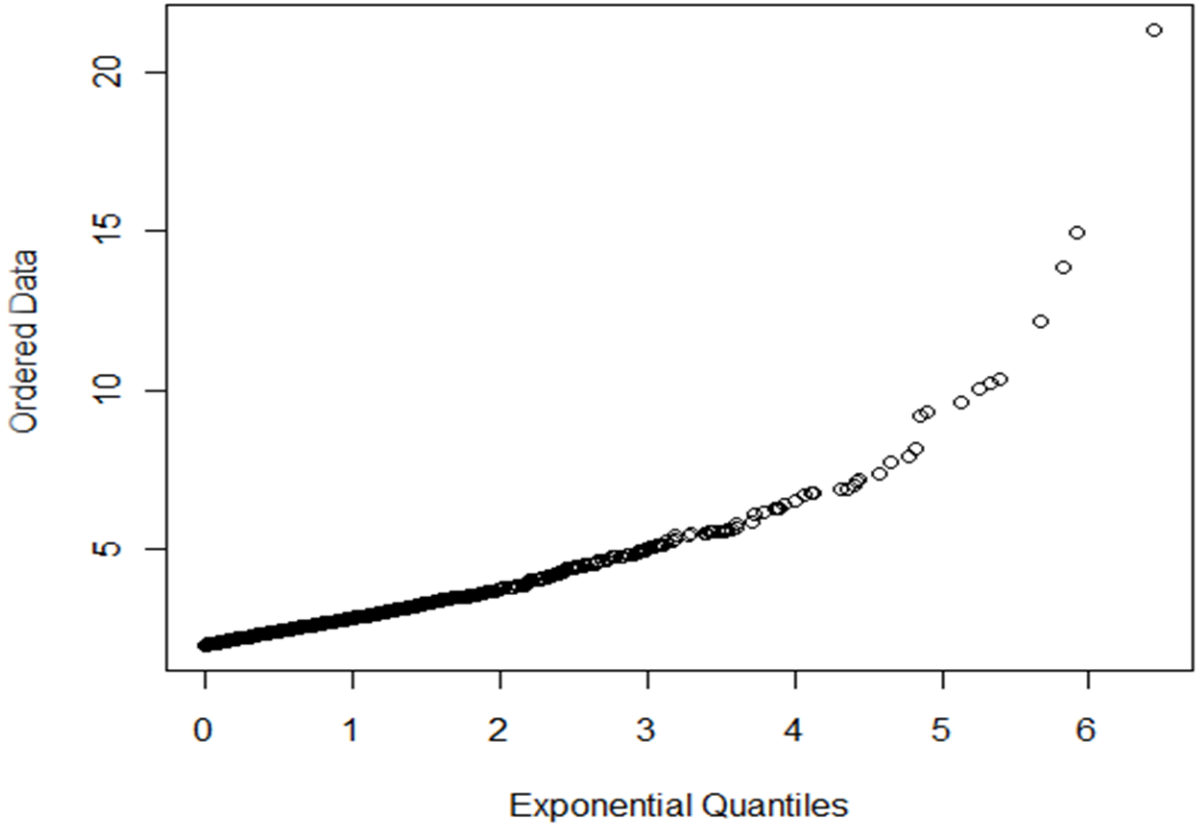
$$Z_k = \sum_{i=1}^{N_k} Y_{k,i}$$

Application of POT model

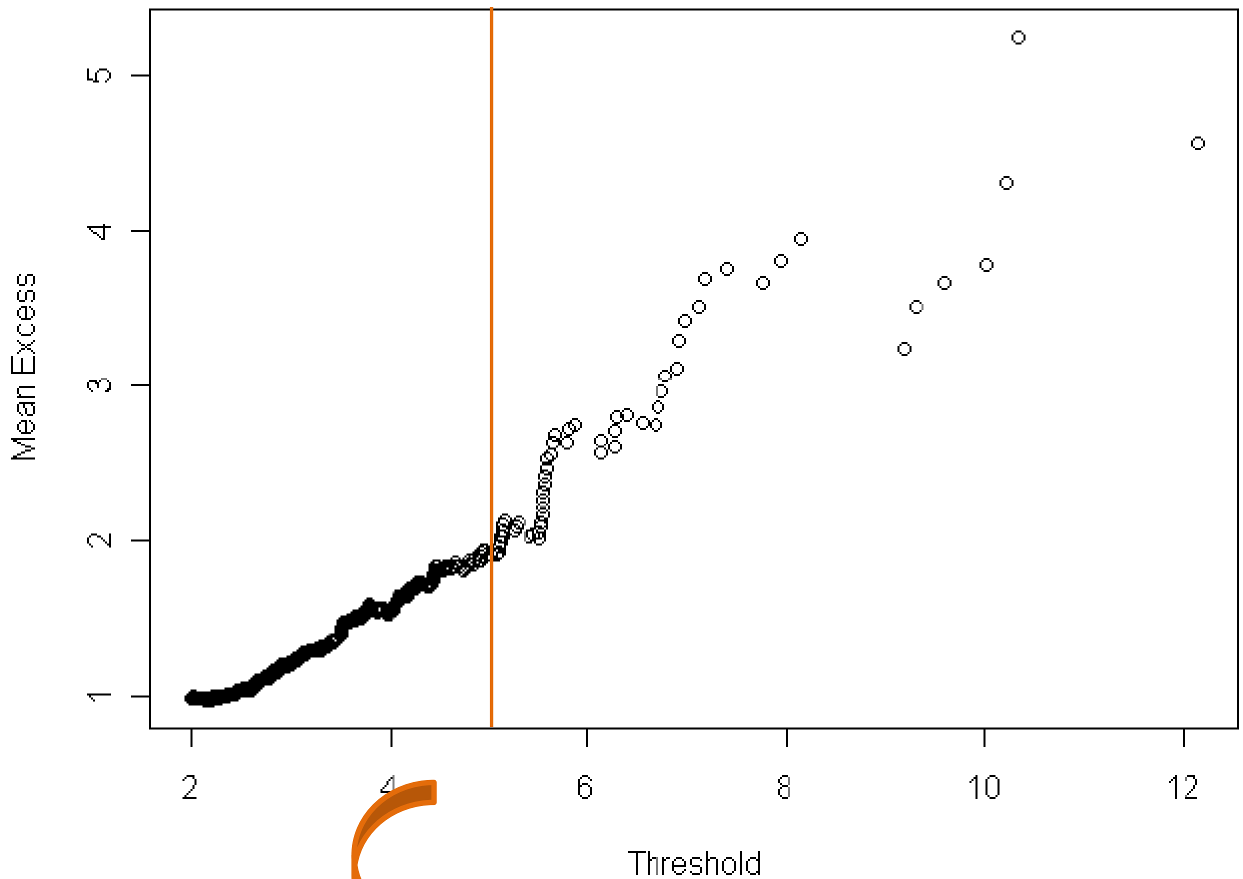
Observed losses during period 2001. - 2008.



Heavy tailed distribution?
QQ plot



Threshold – sample mean excess plot



threshold = 5

Estimated parameters of POT model

$\mu = 2.949$ $\sigma = 0.378$ $\xi = 0.397$

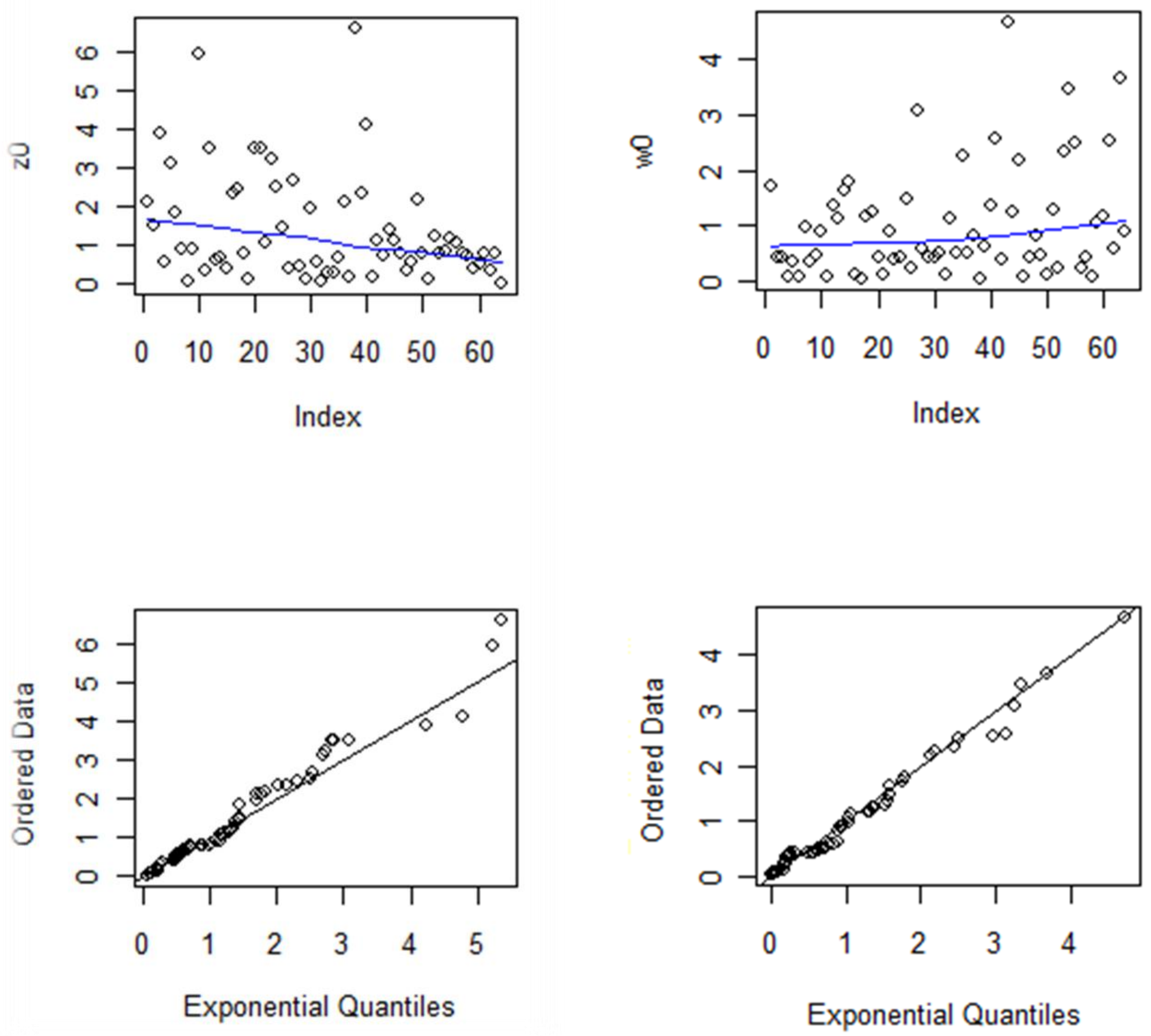
Parameters of generalized Pareto distribution

$$\ln \lambda = -\frac{1}{\xi} \ln \left(1 + \xi \frac{u - \mu}{\sigma}\right)$$

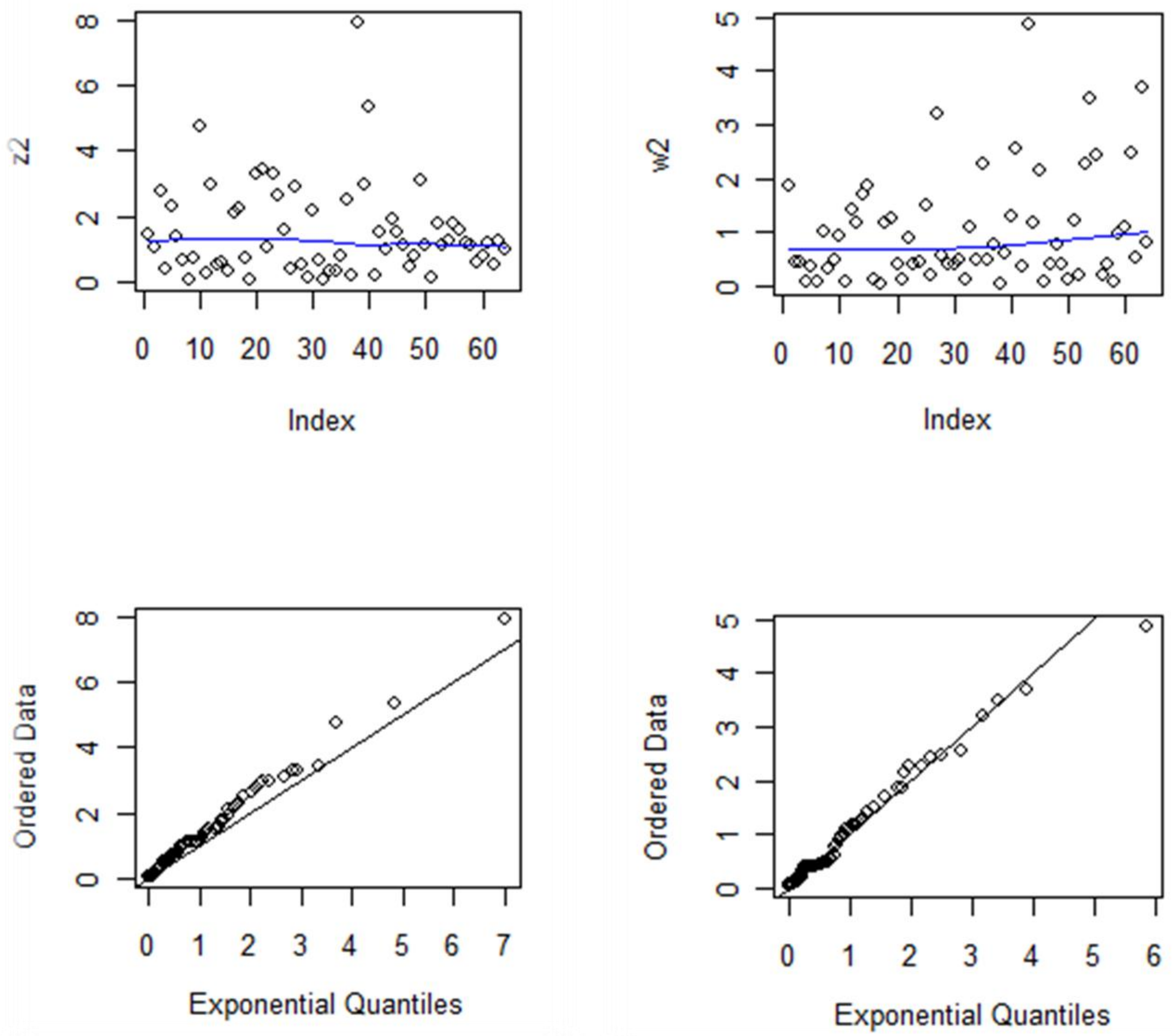
$$\beta = \sigma + \xi(u - \mu),$$

$\xi = 0.397$ $\beta = 1.193.$

Rescaled inter-exceedance times and residuals – Model 0



Rescaled inter-exceedance times and residuals – Model 2



Claim number and loss volume – experience

	Claim number	Loss volume	Average claim
2003	10	58,40	5,84
2004	11	69,42	6,31
2005	15	97,22	6,48
2006	13	98,80	7,60
2007	15	117,46	7,83

Estimated claim number

	Model 0	Model 1	Model 2	Model 3	Model 12
N_1	9,14	13,69	17,24	17,67	20,26
N_2	9,14	15,57	20,29	19,93	25,65
N_3	9,14	17,81	23,69	22,15	32,02

Estimated loss volume

	Model 0	Model 1	Model 2	Model 3	Model 12
Z_1	92,64	132,51	165,55	214,34	159,94
Z_2	92,64	141,17	197,02	294,65	208,70
Z_3	92,64	159,89	238,58	384,00	282,42

Estimated average claim size

	Model 0	Model 1	Model 2	Model 3	Model 12
Z_1/N_1	10,13	9,67	9,60	12,13	7,05
Z_2/N_2	10,13	9,06	9,71	14,78	8,13
Z_3/N_3	10,13	8,97	10,07	17,36	8,82

Model 2 – realistic model for loss volume

Model 3 – safer model for loss volume

"Prediction is very difficult, especially about the future."

N. Bohr

List of references

Embrechts P., Klüppelberg C., Mikosch T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

McNeal A., Saladin T. (2000) Developing Scenarios for Future Extreme Losses using the POT Model. *Extremes and Integrated Risk Management,* RISK books, London.

McNeal A., Saladin T. (1997) The Peaks over Threshold Method for Estimating High Quantiles of Loss Distributions. *Proceedings of XXVIIth International ASTIN Colloquium,* 23-43.

Reiss R.D., Thomas M. (2007) *Statistical Analysis od Extreme Values.* Birkhäuser Verlag, Basel.

Resnick S. (1997) Discussion of the Danish data on large fire insurance losses. *Astin bulletin,* **27**, 139-151.

Thank you for your attention

branka.skelin@agramlife.hr
branka.skelin@gmail.com