Approaching risks using the Extreme Value Theory

- An Application on MTPL -

Pomeriggio Attuariale del 4 Dicembre 2014 - Torino

Valerio Amante - Department of Statistical Science, La Sapienza.

valerio.amante@uniroma1.it
amante@studiosavelli.it

Comitato Regionale del Piemonte

SAPIENZA UNIVERSITÀ DI ROMA

Studio Attuariale Savelli

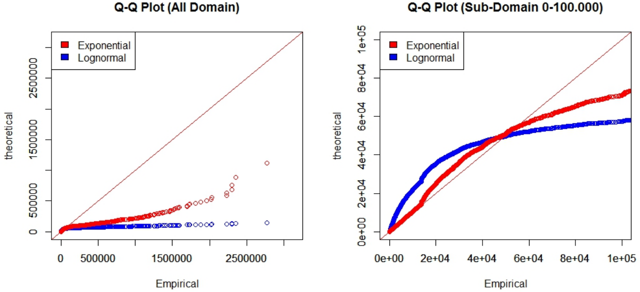
- Considering a real case study refered to a Non-Life insurance company operating in MTPL business, the aim of this work is to model the claim-size;
- In particular, using results of Extreme Value Theory, we will focus on the upper tail of that distribution;
- For this purpose, the way to define large claims becomes a sensible issue. This topic is relevant in the actuarial literature, as connected to the definition of Reinsurance Policy, Solvency Model, and Risk Manage.

The aggregate claim ammount $\widetilde{X}$ is well described by a compound process as the sum of a random number $\widetilde{K}$ of random variabile $\widetilde{Z}_i$ (i.i.d), with $\widetilde{Z}$ and $\widetilde{K}$ indipendent:

$$\widetilde{X} = \sum_{i=1}^{\widetilde{K}} \widetilde{Z}_i$$

⊕ Daykin, Pentikainen, Pesonen (1994)

Generally, and also in our case, the calibration of the unique domain of claim-size doesn't give an acceptable fit using traditional distribution



Separate attritional and large claims

⨁Swiss Solvency Test Approach

$$\widetilde{X} = \widetilde{X}_{Att} + \widetilde{X}_{Lar} = \sum_{i=1}^{\widetilde{K}_{Att}} \widetilde{Z}_{Att,i} + \sum_{i=1}^{\widetilde{K}_{Lar}} \widetilde{Z}_{Lar,i}$$

- $\widetilde{X}_{Att}$ and $\widetilde{X}_{Lar}$ indipendent;
- $\widetilde{X}$ obtained by convolution of $\widetilde{X}_{Att}$ and $\widetilde{X}_{Lar}$

Make a mixture of different distribution

⨁Savelli, Clemente, Zappa (2014)

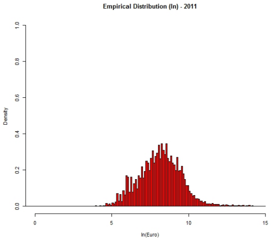
$$\widetilde{X} = \sum_{i=1}^{\widetilde{K}} \widetilde{Z}_i$$

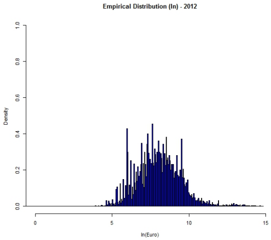
- $\widetilde{Z}$ obtained by mixture of distribution;

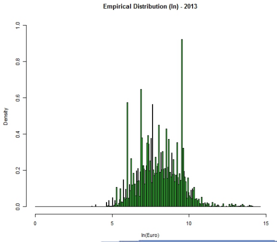
In this context, we use tools of Extreme Value Theory for the identification of Cut-off and calibration of claim-size distribution over this threshold.

Insurance Claims Dataset - (MTPL LoB) - No Card

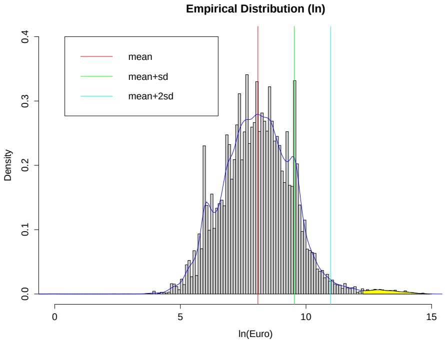
	2011	2012	2013
Mean	14.451	14.814	12.356
Standard Deviation	83.288	87.713	68.252
CoV	5,8	5,9	5,5
Skewness	16,6	15,8	15,7
Kurtosis	365,6	325,6	317,6
Percentile 50%	3.544	2.963	3.000
Percentile 75%	8.443	7.590	8.800
Percentile 95%	31.207	30.360	27.400
Percentile 99.5%	500.643	526.423	440.585
Max	3.127.080	2.766.371	2.300.000
Num. Oss.	14.539	12.189	10.084

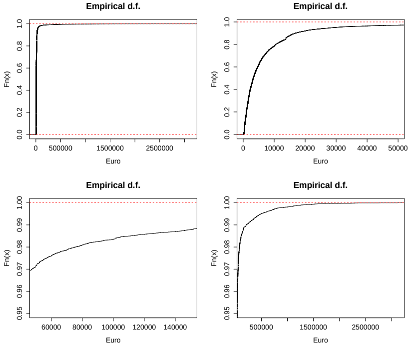






Insurance Claims Dataset - (MTPL LoB)





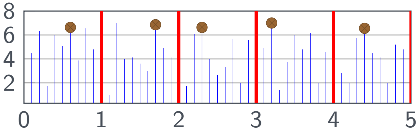
Main Characteristics Empirical Distribution

N.Obs	36.812	Skewness	16,36	Median	3.150	99^{th} Perc.	208.472
Mean	13.997	Kurtosis	353,40	75^{th} Perc.	8.235	Min	0,5
St. Dev.	81.019	10^{th} Perc.	500	95^{th} Perc.	29.266	Max	3.127.080
CoV	5,79	25^{st} Perc.	1.251				

Extreme Value Theory: Parametric Approach

⊕Embrechts, Klüppelberg, Mikosh (1997)

Block Maxima Method (GEV)



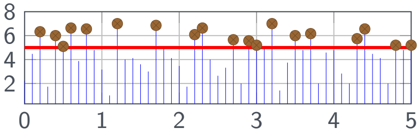
Fisher and Tippett (1928), Gnedenko (1943)

Let (X_n) be a sequence of i.i.d. r.v. If there are constants $c_n > 0$, $d_n \in \mathbb{R}$ and some non-degenerate df H such that $(M_n - d_n)c_n^{-1} \xrightarrow{d} H$ then H belongs to one of:

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

Fréchet	$\xi > 0$
Gumbel	$\xi = 0$
Weibull	$\xi < 0$

Threshold Exceedances Method (GPD)



Pickands (1975), Balkema and de Hann (1974)

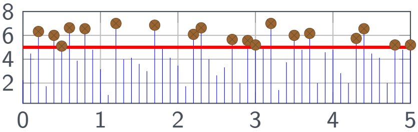
For a large class of underlying distribution functions F the conditional excess distribution function F_u, for u large, is well aproximated by

$$F_u(y) = P(X \leq u + y \mid X > u) \approx G_{\xi,\beta(u)}(y) \quad u \to \infty$$

where $G_{\xi,\beta}(x)$ is the generalized Pareto distribution:

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

Threshold Exceedances Method (GPD)



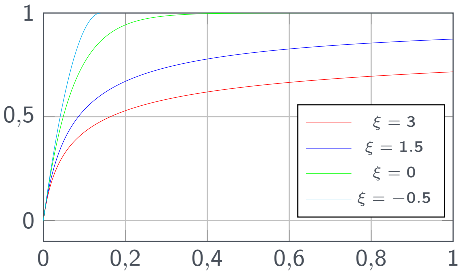
Generalized Pareto distribution (GPD)

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

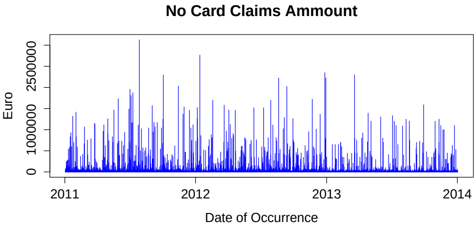
where β scale parameter, ξ shape parameter:

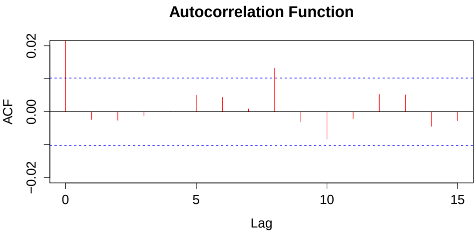
- $x \geq 0$ when $\xi \geq 0$;
- $0 \leq x \leq -\frac{\beta}{\xi}$ when $\xi < 0$;
- if $\xi > 0$ GPD is heavy-tailed and $E(X^k) < \infty$ for $k < \frac{1}{\xi}$

GPD for different shape value



Verify the hypothesis of independence





The basic assumption (more restrictive in financial times series) in EVT is that data be iid.

Casuality Test

Difference Sign

$$\frac{S - \frac{1}{2}(n-1)}{\sqrt{\frac{n+1}{12}}} \sim N(0,1)$$

H0=acc. (0.1537)

Turning Point

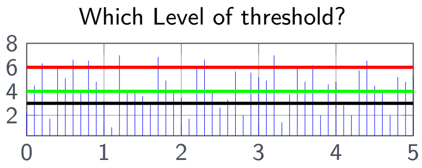
$$\frac{T - \frac{2}{3}(n-2)}{\sqrt{\frac{16n-29}{90}}} \sim N(0,1)$$

H0=acc. (1.618)

Def: Sample Autocorrelation function

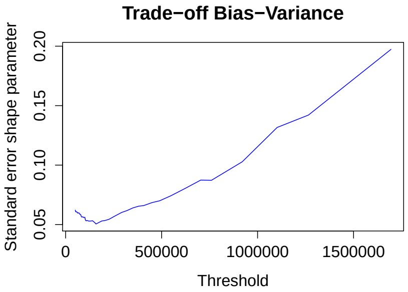
$$\hat{\rho}(h) = \frac{\sum_{t=1}^{n-h}(x_t - \bar{x})(x_{t+h} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

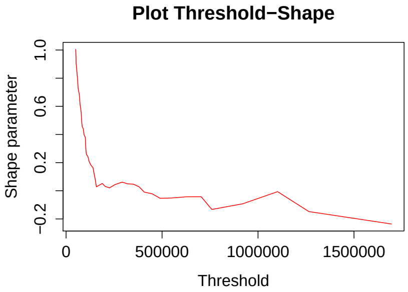
The challenge: Setting the Cut-off



The determination of a cut-off is not easy, since it involves a trade-off between bias and variance:

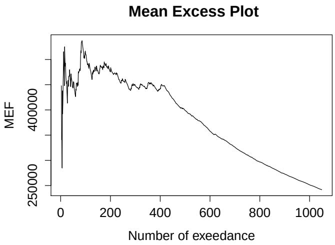
- if the number of exceedance is set too high, many data are included in the tail of the distribution, even if not extreme, yielding biased parameter estimates;
- on the other hand, low values of numbers of exceedance give fewer observations of extreme data, resulting in inefficient estimates of the parameters, with huge standard errors.

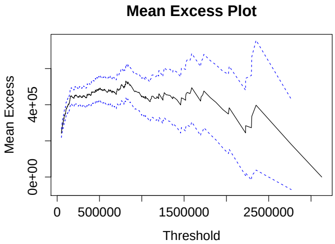




The challenge: Setting the Cut-off

⊕McNeil, Frey, Embrechts (2005)





Excess over higher thresholds

Let F be a distribution with endpoint x_F and for some high threshold u we have $F_u(x) = G_{\xi,\beta}(x)$. From this we can infer model for higher threshold:

$$F_v(x) = G_{\xi,\beta+\xi(v-u)}(x) \qquad e(v) = \frac{\xi v}{1-\xi} + \frac{\beta - \xi u}{1-\xi}$$

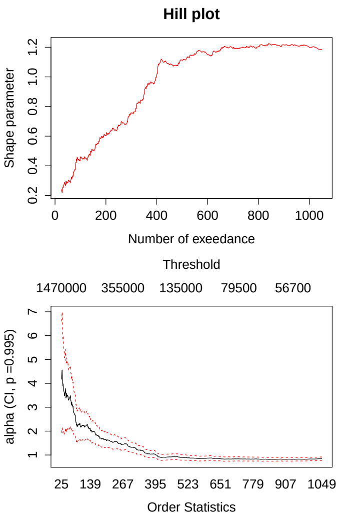
The function remains a GDP with the same shape but a scaling that grows linearity with the threshold v. The linearity of the mean is used as a diagnostic:

$$e_n(u) = \frac{\sum_{i=1}^{n}(X_i - u)I_{\{X_i > u\}}}{\sum_{i=1}^{n} I_{\{X_i > u\}}}$$

- Linear upward trend indicates a GDP with shape $\xi > 0$
- A plot tending toward horizontal indicates a GDP with shape $\xi = 0$
- A Linear downward trend indicates a GDP with shape $\xi < 0$

The challenge: Setting the Cut-off

McNeil, Frey, Embrechts (2005)



Hill estimator Plot

The inverse of the average log-exceedance above the threshold $ln\ X_{k,n}$

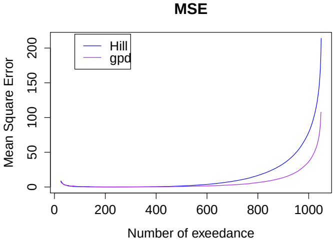
$$\hat{\alpha}_{k,n} = \left[\frac{\sum_{j=1}^{k} (\ln X_{j,n} - \ln X_{k,n})}{k} \right]^{-1}$$

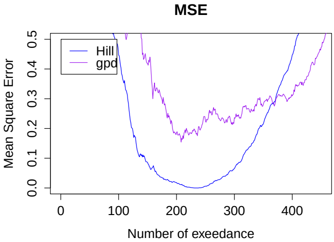
It's a well known different approach from parametric via to estimate tail index, that have good asymptotic properties

- Weak consistency - $\hat{\alpha}_{k,n} \xrightarrow{p} \alpha$ if $k \to \infty$ and $k/n \to 0$ for $n \to \infty$
- Strong consistency - $\hat{\alpha}_{k,n} \xrightarrow{a.s.} \alpha$ if $k/n \to 0$ and $k/\ln\ln n \to \infty$ for $n \to \infty$
- Asymptotic normality - $\sqrt{k}(\hat{\alpha}_{k,n} - \alpha) \xrightarrow{d} N(0, \alpha^2)$

A challenge: Setting the Cut-off

Longin, Solnik (2001)





Monte Carlo Simulation and minimization of MSE

The procedure can be devided in three steps:

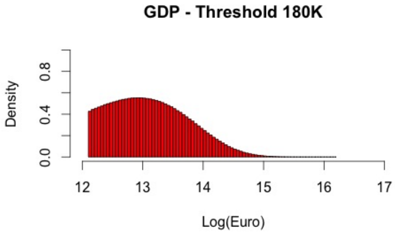
1. First we simulate S observations containing claims ammount from a distribution $F \in MDA(H_\xi)$
   The tail index ξ is related to ϕ by $\xi = f(\phi)$
2. For different numbers n of return exceedances, we obtain a tail index estimate ξ;
3. Calculate the MSE (via the following formula) and choose k that minimize it:

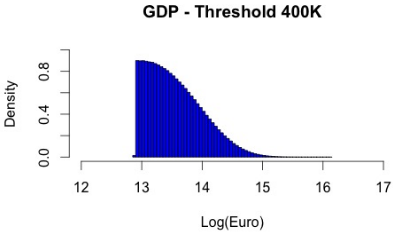
$$MSE(\hat{\xi}) = E[(\hat{\xi} - \xi)]^2 + Var(\xi)$$

$$= (\bar{\xi} - \xi)^2 + \frac{1}{S}\sum_{s=1}^{S}(\hat{\xi} - \xi)^2$$

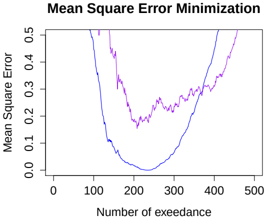
with $\bar{\xi}$ the mean of S simulated observation.

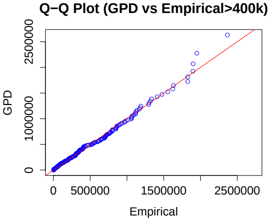
	Threshold 180K	Threshold 400K
Mean	622.348	859.117
SD	465.042	472.392
CoV	0,75	0,55
Skewness	2,3	2,2
Kurtosis	11,9	10,5



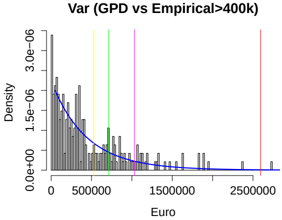


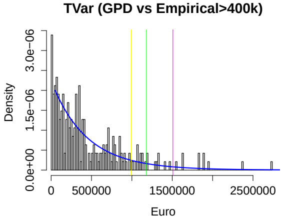
Fitting GDP to the Data



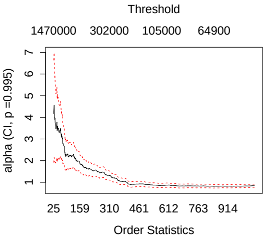


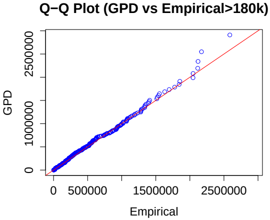

```
$threshold
[1] 4e+05
$n.exceed
[1] 236
$method
[1] "ml"
$par.ests
          xi         beta
2.779602e-02 4.464518e+05
$riskmeasures
         p  quantile     sfall
[1,] 0.250  528951.1  991854.1
[2,] 0.500  712457.2 1180606.8
[3,] 0.750 1030992.8 1508249.5
[4,] 0.990 2593374.5 3115300.8
$TVAR/VAR
1.875134
1.657092
1.462910
1.201254
```



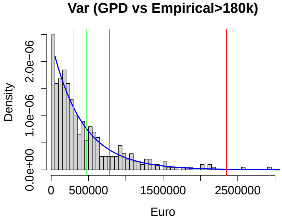


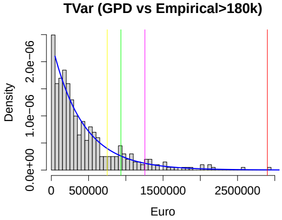
Fitting GDP to the Data






```
$threshold
[1] 180000
$n.exceed
[1] 400
$method
[1] "ml"
$par.ests
           xi         beta
4.768410e-02 4.211489e+05
$riskmeasures
         p  quantile     sfall
[1,] 0.250  301991.8  750336.6
[2,] 0.500  476796.0  933893.6
[3,] 0.750  783565.6 1256023.7
[4,] 0.990 2348890.6 2899727.2
$TVAR/VAR
2.484626
1.958686
1.602959
1.234509
```





- The choice of cut-off is a problem not easy to solve. In our case-study based on real data, every applied methodology identifies different thresholds.
- The results of every used methodology depend on distributional assumptions, mathematical properties associated with the model applied, reading of graphical tools. The attribution of greater credibility to a method rather than another is related to the awareness of the evaluator regarding the reasonableness of the tools used.
- Regarding future researches, my goal is to analyze the multivariate approach (with particular reference to Flooding, Hail and Hurrican losses)

Main references

- Embrechts P., Klüppelberg C., Mikosh T. (1997), *Modelling Extremal Events*, Springer.
- McNeil A., Frey R., Embrechts P. (2005), *Quantitative Risk Management*, Princeton.
- de Haan L., Ferreira A. (2006), *Extreme Value Theory*, Springer.
- McNeil A. (1999), *Estimating the Tails of Loss Severity Distributions using Extreme Value Theory*.
- Klugman S., Panjer H., Willmot G. (2012), *Loss Models*, Wiley.
- Wüthrich M., Merz M. (2013), *Financial Modeling, Actuarial Valuation and Solvency in Insurance*, Springer.
- Savelli N., Clemente G.P., Zappa D. (2014), *Modelling and calibration for non-life underwriting risk: from empirical data to risk capital evalutation*, ICA2014.
- Rocco M. (2011), *Extreme Value Theory for finance*, Questioni di Economia e Finanza.
- Charpentier A. (2007), *Modeling and covering catastrophic risks*, AXA Risk College.
- Longin F. and Solnik B.(2001), *Extreme Correlation of International Equity Markets*, The journal Of Finance.
- Daykin C.D., Pentikainen T., Pesonen M. (1994), *Pratical Risk Theory for Actuaries*, Chapman & Hall.