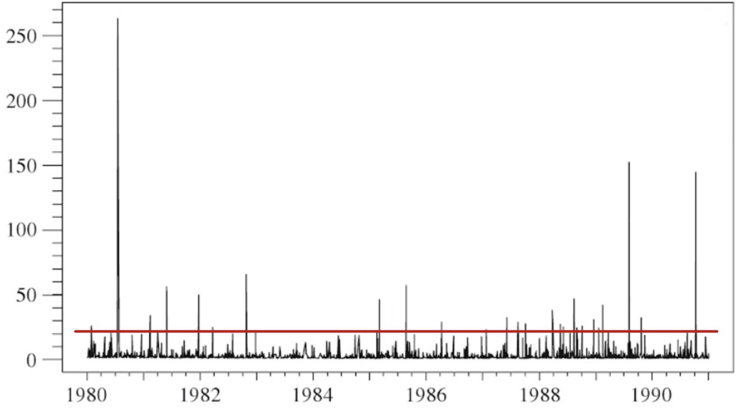
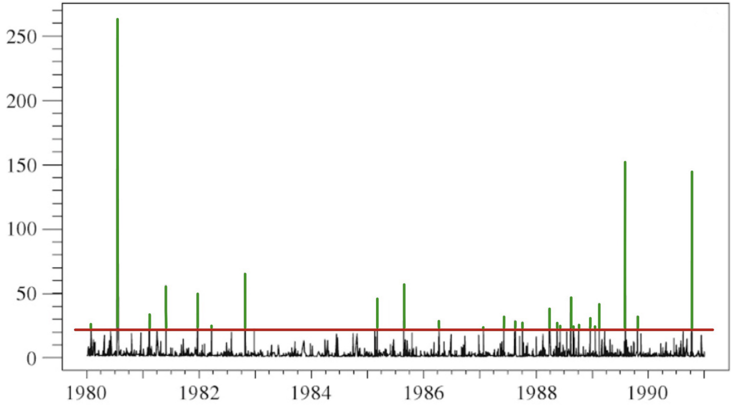
Threshold Exceedances

Moritz Lücke

27. April 2018

1. The GPD Method
   - Estimating ξ and β
   - Estimating the Threshold
   - modelling Tails and Measures of Tail risk
2. The Hill Method
3. Sources


250
200
150
100
50
0
1980
1982
1984
1986
1988
1990


250
200
150
100
50
0
1980
1982
1984
1986
1988
1990

Excess Distribution

Definition

Let X be a rv with df F. The excess distribution over the threshold u has the df

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

for $0 \leqslant x < X_F - u$.

The GPD

Definition

The General Pareto Distribution (GPD) is given by:

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

for $\beta > 0$, $x \geqslant 0$ if $\xi \geqslant 0$ and $0 \leqslant x \leqslant -\frac{\beta}{\xi}$ if $\xi < 0$.

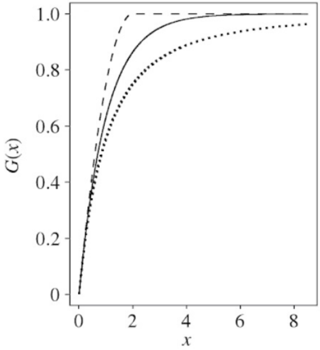


Abbildung: GPD mit $\xi = -0.5; 0; 0.5$ und $\beta = 1$

Theorem

We can find a positive-measurable function $\beta(u)$ *so that*

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

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

Use of the Theorem

We assume that for some high threshold u, we have $F_u(x) = G_{\xi,\beta}(x)$ for $0 \leqslant x < X_F - u$ and some $\xi \in \mathbb{R}$ *and* $\beta > 0$.

Estimating ξ and β

- Given the loss data $X_1, ..., X_n$ from F, a random number N_u will exceed our threshold u.

Estimating ξ and β

- Given the loss data $X_1, ..., X_n$ from F, a random number N_u will exceed our threshold u.
- Relabel these $X'_1, ..., X'_{N_u}$.

Estimating ξ and β

- Given the loss data $X_1, ..., X_n$ from F, a random number N_u will exceed our threshold u.
- Relabel these $X'_1, ..., X'_{N_u}$.
- We write $Y'_j = X'_j - u$.

Estimating ξ and β

- Given the loss data $X_1, ..., X_n$ from F, a random number N_u will exceed our threshold u.
- Relabel these $X'_1, ..., X'_{N_u}$.
- We write $Y_j' = X_j' - u$.
- We can use the log-likelihood method:

$$L(\xi, \beta; Y_1, ..., Y_{N_u}) = \sum_{j=1}^{N_u} \ln(g_{\xi,\beta}(Y_j))$$

$$= -N_u \ln(\beta) - (1 + \frac{1}{\xi}) \sum_{j=1}^{N_u} \ln(1 + \xi \frac{Y_j}{\beta}),$$

Theorem

We can find a positive-measurable function $\beta(u)$ so that

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

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

Definition

The mean excess function of an rv X with finite mean is given by $e(u) = E(X - u|X > u)$.

Definition

The mean excess function of an rv X with finite mean is given by $e(u) = E(X - u|X > u)$.

Theorem

Under the assumption $F_u(x) = G_{\xi,\beta}(x)$, it follows that the mean excess function is linear for all $v > u$

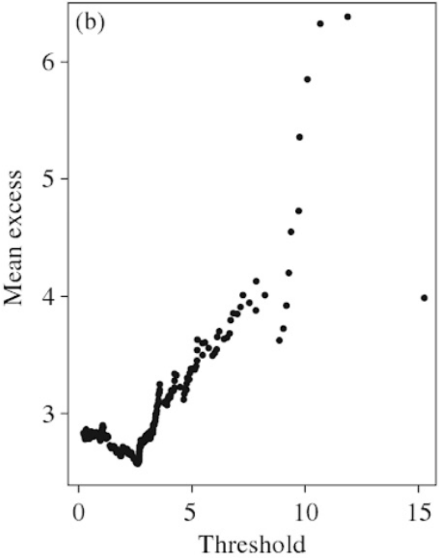
For positive-valued loss data $X_1, ..., X_n$ we estimate the mean excess function with the sample mean excess function given by

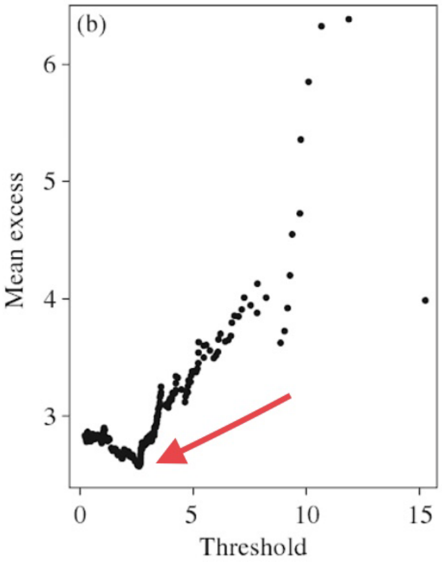
$$e_n(v) = \frac{\sum_{i=1}^{n} (X_i - v) I_{(X_i > v)}}{\sum_{i=1}^{n} I_{(X_i > v)}}.$$

- We construct the mean excess plot $\{(X_{i,n}, e_n(X_{i,n})) : 2 \leqslant i \leqslant n\}$ where X_i denotes the upper ith order statistic.

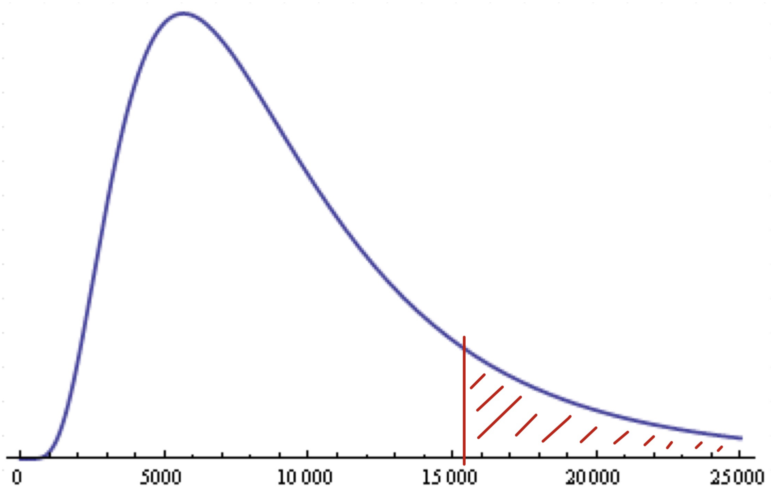
- We construct the mean excess plot $\{(X_{i,n}, e_n(X_{i,n})) : 2 \leqslant i \leqslant n\}$ where X_i denotes the upper ith order statistic.
- If the data support a GDP model over a high threshold, then the plot should become increasingly linear for higher values of v.

- We construct the mean excess plot $\{(X_{i,n}, e_n(X_{i,n})) : 2 \leqslant i \leqslant n\}$ where X_i denotes the upper ith order statistic.
- If the data support a GDP model over a high threshold, then the plot should become increasingly linear for higher values of v.
- By this, we can estimate the needed high of our threshold.


(b)
Mean excess
Threshold
6
5
4
3
0
5
10
15


(b)
Mean excess
6
5
4
3
0
5
10
15
Threshold

modelling Tails and Measures of Tail risk



Our goal is to estimate the tail of a underlying loss distribution F and associated risk measures.

- We have for $x \geqslant u$

$$\begin{aligned}
\bar{F}(x) &= P(X > u)P(X > x|X > u) \\
&= \bar{F}(u)P(X - u > x - u|X > u) \\
&= \bar{F}(u)\bar{F}_u(x - u) \\
&= \bar{F}(u)(1 + \xi\frac{x - u}{\beta})^{-\frac{1}{\xi}}
\end{aligned}$$

- We have for $x \geqslant u$

$$\begin{aligned}\bar{F}(x) &= P(X > u)P(X > x|X > u) \\ &= \bar{F}(u)P(X - u > x - u|X > u) \\ &= \bar{F}(u)\bar{F}_u(x - u) \\ &= \bar{F}(u)(1 + \xi\frac{x - u}{\beta})^{-\frac{1}{\xi}}\end{aligned}$$

- which gives us a formula for tail probabilities, if $F(u)$ is known.

- We have for $x \geqslant u$

$$\begin{aligned}\bar{F}(x) &= P(X > u)P(X > x|X > u)\\ &= \bar{F}(u)P(X - u > x - u|X > u)\\ &= \bar{F}(u)\bar{F}_u(x - u)\\ &= \bar{F}(u)(1 + \xi\frac{x - u}{\beta})^{-\frac{1}{\xi}}\end{aligned}$$

- which gives us a formula for tail probabilities, if $F(u)$ is known.
- If not, we can estimate $\bar{F}(u)$ with the estimator $\frac{N_u}{n}$.

- We have for $x \geqslant u$

$$\begin{aligned}\bar{F}(x) &= P(X > u)P(X > x|X > u)\\ &= \bar{F}(u)P(X-u > x-u|X > u)\\ &= \bar{F}(u)\bar{F}_u(x-u)\\ &= \bar{F}(u)(1+\xi\frac{x-u}{\beta})^{-\frac{1}{\xi}}\end{aligned}$$

- which gives us a formula for tail probabilities, if $F(u)$ is known.
- If not, we can estimate $\bar{F}(u)$ with the estimator $\frac{N_u}{n}$.
- $\Rightarrow \hat{\bar{F}}(x) = \frac{N_u}{n}(1+\hat{\xi}\frac{x-u}{\hat{\beta}})^{-\frac{1}{\hat{\xi}}}$.

- By inverting the formula, we can obtain a high quantile of the underlying distribution, which we can interpret as a *VaR*. For $\alpha \leqslant F(u)$ we have

- By inverting the formula, we can obtain a high quantile of the underlying distribution, which we can interpret as a *VaR*. For $\alpha \leqslant F(u)$ we have

$$VaR_\alpha = q_\alpha(F) = u + \frac{\beta}{\xi}((\frac{1-\alpha}{\bar{F}(u)})^{-\xi} - 1).$$

- By inverting the formula, we can obtain a high quantile of the underlying distribution, which we can interpret as a VaR. For $\alpha \leqslant F(u)$ we have

$$VaR_\alpha = q_\alpha(F) = u + \frac{\beta}{\xi}((\frac{1-\alpha}{\bar{F}(u)})^{-\xi} - 1).$$

- Assuming $\xi < 1$, the associated expected shortfall is given by

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_x(F)dx = \frac{VaR_\alpha}{1-\xi} + \frac{\beta - \xi u}{1-\xi}.$$

The Hill Method

- Alternative to the GPD Method.

The Hill Method

- Alternative to the GPD Method.
- New assumption: $F \in MDA(H_\xi)$, $\xi > 0$.

The Hill Method

- Alternative to the GPD Method.
- New assumption: $F \in MDA(H_\xi)$, $\xi > 0$.
- We can use the Fréchet Theorem
  ($F \in MDA(H_\xi) \Leftrightarrow \bar{F} = x^{-\frac{1}{\xi}} L(x)$ for $\xi > 0$)

The Hill Method

- Alternative to the GPD Method.
- New assumption: $F \in MDA(H_\xi)$, $\xi > 0$.
- We can use the Fréchet Theorem ($F \in MDA(H_\xi) \Leftrightarrow \bar{F} = x^{-\frac{1}{\xi}} L(x)$ for $\xi > 0$)
- $\Rightarrow \bar{F} = x^{-\alpha} L(x)$ for a function $L \in R_0$ and $\alpha = \frac{1}{\xi} > 0$.

Estimating α

- Given the data $X_1, \ldots, X_n$, we first build the orderstatistic $X_{n,n} \leqslant \ldots \leqslant X_{2,n} \leqslant X_{1,n}$

Estimating α

- Given the data $X_1, ..., X_n$, we first build the orderstatistic $X_{n,n} \leqslant ... \leqslant X_{2,n} \leqslant X_{1,n}$
- the hill estimator is then given by

Estimating α

- Given the data $X_1, \ldots, X_n$, we first build the orderstatistic $X_{n,n} \leqslant \ldots \leqslant X_{2,n} \leqslant X_{1,n}$
- the hill estimator is then given by

$$\hat{\alpha}^H_{k,n} = (\frac{1}{k}\sum_{j=1}^{k} ln(X_{j,n}) - ln(X_{k,n}))^{-1}$$

for $2 \leqslant k \leqslant n$.

Estimating α

- Given the data $X_1, ..., X_n$, we first build the orderstatistic $X_{n,n} \leqslant ... \leqslant X_{2,n} \leqslant X_{1,n}$
- the hill estimator is then given by

$$\hat{\alpha}^H_{k,n} = (\frac{1}{k}\sum_{j=1}^{k} ln(X_{j,n}) - ln(X_{k,n}))^{-1}$$

for $2 \leqslant k \leqslant n$.

Estimating α

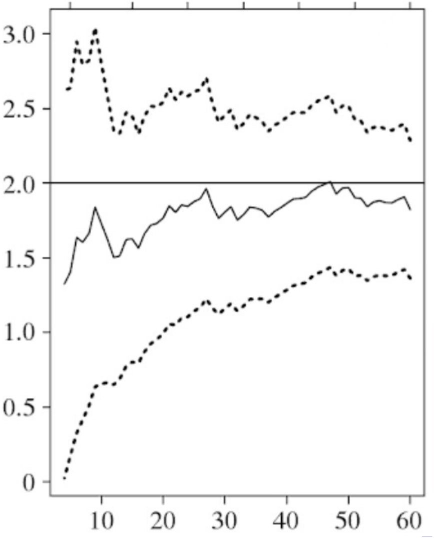
- Given the data $X_1, ..., X_n$, we first build the orderstatistic $X_{n,n} \leqslant ... \leqslant X_{2,n} \leqslant X_{1,n}$
- the hill estimator is then given by

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

for $2 \leqslant k \leqslant n$.

- The strategy is to plot Hill estimates for various values of k. This gives the Hill plot $((k, \hat{\alpha}_{k,n}^{H}) : k = 2, ..., n)$. We hope to find a stable region in the Hill plot.

Estimating α



hill based tail estimates

- We assume a tail of the form $\bar{F}(x) = Cx^{-\alpha}$, $x \geqslant u > 0$ for some high threshold u.

hill based tail estimates

- We assume a tail of the form $\bar{F}(x) = Cx^{-\alpha}$, $x \geqslant u > 0$ for some high threshold u.
- We estimate α by $\hat{\alpha}_{k,n}^{(H)}$ and u by $X_{k,n}$

estimating C

- $\bar{F}(u) = Cu^{-\hat{\alpha}_{k,n}^{(H)}}$

estimating C

- $\bar{F}(u) = Cu^{-\hat{\alpha}_{k,n}^{(H)}} \Leftrightarrow C = u^{\hat{\alpha}_{k,n}^{(H)}} \bar{F}(u)$

estimating C

- $\bar{F}(u) = Cu^{-\hat{\alpha}_{k,n}^{(H)}} \Leftrightarrow C = u^{\hat{\alpha}_{k,n}^{(H)}} \bar{F}(u)$
- The empirical estimator for $\bar{F}(u)$ is $\frac{k}{n}$.

estimating C

- $\bar{F}(u) = Cu^{-\hat{\alpha}_{k,n}^{(H)}} \Leftrightarrow C = u^{\hat{\alpha}_{k,n}^{(H)}} \bar{F}(u)$
- The empirical estimator for $\bar{F}(u)$ is $\frac{k}{n}$.
- We get the Hill tail estimator

$$\hat{\bar{F}}(x) = \frac{k}{n}\left(\frac{x}{X_{k,n}}\right)^{-\hat{\alpha}_{k,n}^{(H)}}, x \leqslant X_{k,n}.$$

Sources

[1] A.McNeil R.Frey P.Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton Series in Finance, 2015.