Hill estimators for the tail distribution and the quantile

Hill estimators for the tail distribution and the quantile

Thus the following Hill estimator is consistent:

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

Hill estimators for the tail distribution and the quantile

Thus the following Hill estimator is consistent:

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

How to choose a suitable k for a given sample size n?

Hill estimators for the tail distribution and the quantile

Thus the following Hill estimator is consistent:

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

How to choose a suitable k for a given sample size n?
If k too small, then the estimator has a high variance.

Hill estimators for the tail distribution and the quantile

Thus the following Hill estimator is consistent:

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

How to choose a suitable k for a given sample size n?
If k too small, then the estimator has a high variance.
If k too large, than the estimator is based on central values of the sample distribution, and is therefore biased.

Hill estimators for the tail distribution and the quantile

Thus the following Hill estimator is consistent:

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

How to choose a suitable k for a given sample size n?
If k too small, then the estimator has a high variance.
If k too large, than the estimator is based on central values of the sample distribution, and is therefore biased.

Grafical inspection of the Hill plots: $\left\{ \left(k, \hat{\alpha}_{k,n}^{(H)} \right) : k = 2, \ldots, n \right\}$

Hill estimators for the tail distribution and the quantile

Thus the following Hill estimator is consistent:

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

How to choose a suitable k for a given sample size n?
If k too small, then the estimator has a high variance.
If k too large, than the estimator is based on central values of the sample distribution, and is therefore biased.

Grafical inspection of the Hill plots: $\left\{ \left(k, \hat{\alpha}_{k,n}^{(H)} \right) : k = 2, \ldots, n \right\}$

Given an estimator $\hat{\alpha}_{k,n}^{(H)}$ of α we get tail and quantile estimators as follows:

$$\hat{\bar{F}}(x) = \frac{k}{n} \left(\frac{x}{x_{k,n}} \right)^{-\hat{\alpha}_{k,n}^{(H)}} \text{ and } \hat{q}_p = \hat{F}^{\leftarrow}(p) = \left(\frac{n}{k}(1-p) \right)^{-1/\hat{\alpha}_{k,n}^{(H)}} x_{k,n}.$$

The POT method (Peaks over Threshold)

The POT method (Peaks over Threshold)

Definition: (The generalized Pareto distribution (GPD))
The standard GPD denoted by G_γ:

$$G_\gamma(x) = \begin{cases} 1 - (1+\gamma x)^{-1/\gamma} & \text{für } \gamma \neq 0 \\ 1 - \exp\{-x\} & \text{für } \gamma = 0 \end{cases}$$

where $x \in D(\gamma)$

$$D(\gamma) = \begin{cases} 0 \leq x < \infty & \text{für } \gamma \geq 0 \\ 0 \leq x \leq -1/\gamma & \text{für } \gamma < 0 \end{cases}$$

The POT method (Peaks over Threshold)

Definition: (The generalized Pareto distribution (GPD))
The standard GPD denoted by G_γ:

$$G_\gamma(x) = \begin{cases} 1 - (1+\gamma x)^{-1/\gamma} & \text{für } \gamma \neq 0 \\ 1 - \exp\{-x\} & \text{für } \gamma = 0 \end{cases}$$

where $x \in D(\gamma)$

$$D(\gamma) = \begin{cases} 0 \leq x < \infty & \text{für } \gamma \geq 0 \\ 0 \leq x \leq -1/\gamma & \text{für } \gamma < 0 \end{cases}$$

Notice that $G_0 = \lim_{\gamma \to 0} G_\gamma$.

The POT method (Peaks over Threshold)

Definition: (The generalized Pareto distribution (GPD))
The standard GPD denoted by G_γ:

$$G_\gamma(x) = \begin{cases} 1 - (1+\gamma x)^{-1/\gamma} & \text{für } \gamma \neq 0 \\ 1 - \exp\{-x\} & \text{für } \gamma = 0 \end{cases}$$

where $x \in D(\gamma)$

$$D(\gamma) = \begin{cases} 0 \leq x < \infty & \text{für } \gamma \geq 0 \\ 0 \leq x \leq -1/\gamma & \text{für } \gamma < 0 \end{cases}$$

Notice that $G_0 = \lim_{\gamma \to 0} G_\gamma$.

Let $\nu \in \mathbb{R}$ and $\beta > 0$. The GPD with parameters γ, ν, β is given by the following distribution function

$$G_{\gamma,\nu,\beta} = 1 - (1 + \gamma \frac{x-\nu}{\beta})^{-1/\gamma}$$

where $x \in D(\gamma, \nu, \beta)$ and

$$D(\gamma,\nu,\beta) = \begin{cases} \nu \leq x < \infty & \text{für } \gamma \geq 0 \\ \nu \leq x \leq \nu - \beta/\gamma & \text{für } \gamma < 0 \end{cases}$$

Characteristions of $MDA(H_\gamma)$

Characteristions of $MDA(H_\gamma)$

Theorem: Let $\gamma \in \mathbb{R}$. The following statements are equiavlent:

(i) $F \in MDA(H_\gamma)$

(ii) There exists a positive measurable function $a(\cdot)$, such that for $x \in D(\gamma)$
$$\lim_{u \uparrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \bar{G}_\gamma(x) \text{ holds.}$$

Characteristions of $MDA(H_\gamma)$

Theorem: Let $\gamma \in \mathbb{R}$. The following statements are equiavlent:

(i) $F \in MDA(H_\gamma)$

(ii) There exists a positive measurable function $a(\cdot)$, such that for $x \in D(\gamma)$

$$\lim_{u \uparrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \bar{G}_\gamma(x) \text{ holds.}$$

Definition:(Excess distribution)

Let X be a r.v. with distribution function F and let x_F be the right tail of this distribution. For $u < x_F$ the function F_u given as

$$F_u(x) := P(X - u \leq x | X > u), x \geq 0$$

ic called *excess distribution function over the threshold* u.

Characteristions of $MDA(H_\gamma)$

Theorem: Let $\gamma \in \mathbb{R}$. The following statements are equiavlent:

(i) $F \in MDA(H_\gamma)$

(ii) There exists a positive measurable function $a(\cdot)$, such that for $x \in D(\gamma)$

$$\lim_{u \uparrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \bar{G}_\gamma(x) \text{ holds.}$$

Definition:(Excess distribution)
Let X be a r.v. with distribution function F and let x_F be the right tail of this distribution. For $u < x_F$ the function F_u given as

$$F_u(x) := P(X - u \le x | X > u), x \ge 0$$

ic called *excess distribution function over the threshold u*.

Theorem: Let $\gamma \in \mathbb{R}$. The following statements are equivalent:

(i) $F \in MDA(H_\gamma)$

(ii) There exists a positive measurable function $\beta(\cdot)$, such that

$$\lim_{u \uparrow x_F} \sup_{x \in (0, x_F - u)} |F_u(x) - G_{\gamma,0,\beta(u)}(x)| = 0 \text{ holds.}$$

POT: estimators for the tail and the quantile of the excess distribution

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \dots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

- Choose a threshold u (high enough, by means of suitable statistical approaches) and compute

$$N_u := |\{i \in \{1, 2, \ldots, n\} : X_i > u\}|$$

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

- Choose a threshold u (high enough, by means of suitable statistical approaches) and compute

$$N_u := |\{i \in \{1, 2, \ldots, n\} : X_i > u\}|$$

- Let $Y_1, Y_2, \ldots, Y_{N_u}$ be the exceedances. Determine $\hat{\beta}$ and $\hat{\gamma}$, such that the following holds:

$$\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma}, 0, \widehat{\beta(u)}}(y),$$

  where $\bar{F}_u(y) = P(X - u > y | X > u)$.

POT: estimators for the tail and the quantile of the excess distribution

Let $X_1, \ldots, X_n$ i.i.d. r.v. with distribution function $F \in MDA(H_\gamma)$ for $\gamma \in \mathbb{R}$.

- Choose a threshold u (high enough, by means of suitable statistical approaches) and compute

$$N_u := |\{i \in \{1, 2, \ldots, n\} \colon X_i > u\}|$$

- Let $Y_1, Y_2, \ldots, Y_{N_u}$ be the exceedances. Determine $\hat{\beta}$ and $\hat{\gamma}$, such that the following holds:

$$\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\widehat{\beta(u)}}(y),$$

  where $\bar{F}_u(y) = P(X - u > y | X > u)$.

- Use N_u and $\bar{F}_u \approx \bar{G}_{\hat{\gamma},0,\widehat{\beta(u)}}$ to obtain estimators for the tail and the quantile of F

$$\bar{F}\widehat{(u+y)} = \frac{N_u}{n}\left(1 + \hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}} \text{ and } \hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\gamma}} - 1\right)$$

POT contd.: How to choose a suitable threshold u?

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

Basic idea: inspect the plot of the *empirical mean excess function* and choose a threshold u_0, such that the empirical mean excess function is approximately linear for $u > u_0$.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

Basic idea: inspect the plot of the *empirical mean excess function* and choose a threshold u_0, such that the empirical mean excess function is approximately linear for $u > u_0$.

The justification :

- $e_F(u) = \int_0^\infty t dF_u(t) \approx \int_0^\infty t dG_{\gamma,0,\beta(u)}(t) = E(G_{\gamma,0,\beta(u)}) = \frac{\beta(u)}{1-\gamma}$, if $F_u(t) \approx G_{\gamma,0,\beta(u)}(t)$.

POT contd.: How to choose a suitable threshold u?

- Is u too large, then there are only a few observed exceedances and not enough data to estimate $\hat{\beta}$ und $\hat{\gamma}$.
- Is u too small, then the approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma},0,\hat{\beta}(u)}(y)$ is not good.

Basic idea: inspect the plot of the *empirical mean excess function* and choose a threshold u_0, such that the empirical mean excess function is approximately linear for $u > u_0$.

The justification :

- $e_F(u) = \int_0^\infty t dF_u(t) \approx \int_0^\infty t dG_{\gamma,0,\beta(u)}(t) = E(G_{\gamma,0,\beta(u)}) = \frac{\beta(u)}{1-\gamma}$, if $F_u(t) \approx G_{\gamma,0,\beta(u)}(t)$.
- If $\bar{F}_u(x) \approx \bar{G}_{\gamma,0,\beta}(x)$ then $\forall v \geq u$ the approximation $\bar{F}_v(x) \approx \bar{G}_{\gamma,0,\beta+\gamma(v-u)}(x)$ holds.

POT (contd.): How to choose a suitable threshold u?

POT (contd.): How to choose a suitable threshold u?

Definition: The empirical mean excess function:
Let $x_1, x_2, \ldots, x_n$ be a sample of i.i.d r.v. Let $N_u = |\{i\colon 1 \le i \le n, x_i > u\}|$ be the number of the sample points which exceed u. The empirical mean excess function $e_n(u)$ is defined as:

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

POT (contd.): How to choose a suitable threshold u?

Definition: The empirical mean excess function:
Let $x_1, x_2, \ldots, x_n$ be a sample of i.i.d r.v. Let $N_u = |\{i\colon 1 \le i \le n, x_i > u\}|$ be the number of the sample points which exceed u. The empirical mean excess function $e_n(u)$ is defined as:

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

Consider the plot of the (interpolation of the) empirical mean excess function: $(x_{k,n}, e_n(x_{k,n}))$, $k = 1, 2, \ldots, n-1$. If this plot is approximately linear around some $x_{k,n}$, then $u := x_{k,n}$ might be a good choice for the threshold value.

POT (contd.): Estimation of the paramters γ and β

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

The likelihood function $L(\gamma, \beta, Y_1, \ldots, Y_{N_u})$ is the conditional probability that $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ under the condition that the observed exceedances are $Y_1, Y_2, \ldots, Y_{N_u}$.

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

The likelihood function $L(\gamma, \beta, Y_1, \ldots, Y_{N_u})$ is the conditional probability that $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ under the condition that the observed exceedances are $Y_1, Y_2, \ldots, Y_{N_u}$.

The following holds:

$$\ln L(\gamma, \beta, Y_1, \ldots, Y_{N_u}) = -N_u \ln \beta - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\gamma}{\beta} Y_i\right)$$

where $Y_i \geq 0$ for $\gamma > 0$ and $0 \leq Y_i \leq -\beta/\gamma$ for $\gamma < 0$.

POT (contd.): Estimation of the paramters γ and β

Let u be a given threshold and let $Y_1, Y_2, \ldots, Y_{N_u}$ be the observed data from the sample which exceed u.

The likelihood function $L(\gamma, \beta, Y_1, \ldots, Y_{N_u})$ is the conditional probability that $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ under the condition that the observed exceedances are $Y_1, Y_2, \ldots, Y_{N_u}$.

The following holds:

$$\ln L(\gamma, \beta, Y_1, \ldots, Y_{N_u}) = -N_u \ln \beta - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{N_u} \ln \left(1 + \frac{\gamma}{\beta} Y_i\right)$$

where $Y_i \geq 0$ for $\gamma > 0$ and $0 \leq Y_i \leq -\beta/\gamma$ for $\gamma < 0$.
(see Daley, Veve-Jones (2003) and Coles (2001))

POT contd.

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

There is an uncertainty related to the more or less arbitrary choice of the threshold u. It can be reduced by

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

There is an uncertainty related to the more or less arbitrary choice of the threshold u. It can be reduced by

- investigating the dependency of the ML-estimator $\hat{\gamma}$ on u.

POT contd.

The maximizers $\hat{\gamma}$ and $\hat{\beta}$ of the log-likelihood function are used as estimators for γ and β (ML-estimators)

The method works well for $\gamma > -1/2$.

The ML-estimators are in this case normally distributed:

$$(\hat{\gamma} - \gamma, \frac{\hat{\beta}}{\beta} - 1) \sim N(0, \Sigma^{-1}/N_u) \text{ where } \Sigma^{-1} = \begin{pmatrix} 1+\gamma & -1 \\ -1 & 2 \end{pmatrix}.$$

There is an uncertainty related to the more or less arbitrary choice of the threshold u. It can be reduced by

- investigating the dependency of the ML-estimator $\hat{\gamma}$ on u.
- visualizing and inspecting the estimated tail distribution

$$\hat{\bar{F}}(u+y) = \frac{N_u}{n}\left(1 + \hat{\gamma}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\gamma}}$$