

Die POT Methode (Peaks over Threshold)

Definition 18 (Die verallgemeinerte Pareto Verteilung (GPD))

Die standard GPD G_γ :

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

wobei $x \in D(\gamma)$

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

oder $G_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \in D(\gamma)$ für $\gamma \neq 0$, und $G_0 = \lim_{\gamma \rightarrow 0} G_\gamma$.

Sei $\nu \in \mathbb{R}$ und $\beta > 0$. Eine GPD ist durch die untenstehende Verteilungsfunktion gegeben

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

wobei $x \in D(\gamma, \nu, \beta)$ und

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

Theorem 15 (Charakterisierung von $MDA(H_\gamma)$)

Sei $\gamma \in \mathbb{R}$. Die untenstehenden Aussagen sind äquivalent:

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

(ii) Es existiert eine positive messbare Funktion $a(\cdot)$, sodass für $x \in D(\gamma)$

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

Definition 19 (Exzess-Verteilung)

Sei X eine ZV mit Verteilungsfunktion F und rechtem Endpunkt x_F .

Für $u < x_F$

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

heißt Exzess-Verteilungsfunktion über die Schwelle u .

Theorem 16 (Eine weitere Charakterisierung von $MDA(H_\gamma)$)

Sei $\gamma \in \mathbb{R}$. Folgende Aussagen sind äquivalent:

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

(ii) Es existiert eine positive messbare Funktion $\beta(\cdot)$, sodass

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

POT: Schätzer für den Tail und das Quantil der Exzess-Verteilung

Seien X_1, \dots, X_n i.i.d. ZV mit Verteilungsfunktion $F \in MDA(H_\gamma)$ für $\gamma \in \mathbb{R}$.

- Wähle eine hohe Schwelle u (unter Verwendung von geeigneten stat. Verfahren) und berechne

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

- Sei Y_1, Y_2, \dots, Y_{N_u} eine Stichprobe von Exzess-Beobachtungen. Bestimme $\hat{\beta}$ und $\hat{\gamma}$, sodass folgendes gilt:

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

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

- Kombiniere die obigen zwei Schritte um folgende Schätzer zu berechnen:

$$\widehat{\bar{F}(u + y)} = \frac{N_u}{n} \left(1 + \hat{\gamma} \frac{y}{\hat{\beta}} \right)^{-1/\hat{\gamma}} \quad (9)$$

$$\hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\hat{\gamma}} - 1 \right) \quad (10)$$

Wie wird eine hohe Schwelle u (POT Methode) gewählt?

- u zu groß: Wenige Beobachtungen für die Schätzung von $\hat{\beta}$ und $\hat{\gamma}$.
- u zu klein: die Approximation $\bar{F}_u(y) \approx \bar{G}_{\hat{\gamma}, \hat{\beta}(u)}(u)$ ist nicht gut genug.

Grundidee: Inspektion des Plots der empirischen durchschnittlichen Exzess-Funktion und Auswahl einer Schwelle u_0 , sodass die empirische durchschnittliche Exzess-Funktion für $u > u_0$ annähernd linear ist.

Begründung: die Durchschnittliche Exzess-Funktion der $GPD_{\gamma,0,\beta}$ ist linear!

Die empirische durchschnittliche Exzess-Funktion

Seien X_1, X_2, \dots, X_n i.i.d ZV.

Sei $N_u = \#\{i: 1 \leq i \leq n, X_i > u\}$ die Anzahl der Überschreitung von u durch X_i

Die empirische durchschnittliche Exzess-Funktion $e_n(u)$:

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

Der Plot der durchschnittlichen Exzess-Funktion: $(X_{k,n}, e_n(X_{k,n}))$ für $k = 1, 2, \dots, n - 1$.

Wenn dieser Plot annähernd linear mit einem positiven Gradienten ist, so wird angenommen, dass die Verteilung einen heavy-tailed Pareto-ähnlichen Tail hat.

Schätzung der Parameter γ und β

Sei u eine gegebene Schwelle und Y_1, Y_2, \dots, Y_{N_u} Beobachtungen der Überschüsse $X_i > u$, $1 \leq i \leq n$.

Die Log-Likelihood Funktion:

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

wobei $Y_i \geq 0$ für $\gamma > 0$ und $0 \leq Y_i \leq -\beta/\gamma$ für $\gamma < 0$.

$L(\gamma, \beta, Y_1, \dots, Y_{N_u})$ ist die bedingte Wahrscheinlichkeit, dass $\bar{F}_u(y) \approx \bar{G}_{\gamma,0,\beta}(y)$ unter der Bedingung, dass die Beobachtungen der Überschüsse Y_1, Y_2, \dots, Y_{N_u} sind.

Für die Ermittlung der Likelihood Funktion siehe Daley, Veve-Jones (2003) and Coles (2001).

Als Schätzer $\hat{\gamma}$ und $\hat{\beta}$ werden jene Werte von γ bzw. β verwendet, die die log-Likelihood Funktion maximieren (ML-Schätzer)

Die Methode funktioniert gut für $\gamma > -1/2$.

Die ML-Schätzer sind in diesem Fall normal verteilt:

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

Um die Unsicherheit über die einigermaßen willkürliche Auswahl von u zu reduzieren, überprüft man wie die ML-Schätzer in Abhängigkeit von u variieren.

Weiters wird der Schätzer

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

grafisch dargestellt und inspiziert.

Berechnung von Risikomaßen VaR und CVaR mit Hilfe der POT Methode

Seien X_1, X_2, \dots, X_n Beobachtungen von i.i.d. ZVen mit unbekannter Verteilungsfunktion F . Direkt aus der POT Methode erhält man folgende Schätzer für die Randverteilung und das Quantil $q_p = VaR_p(F)$ von F

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

$$\hat{q}_p = u + \frac{\hat{\beta}}{\hat{\gamma}} \left(\left(\frac{n}{N_u} (1 - p) \right)^{-\hat{\gamma}} - 1 \right)$$

Für $0 < \hat{\gamma} < 1$ zeigen wir, dass $\widehat{CVaR}_p(F) = \hat{q}_p + \frac{\hat{\beta} + \hat{\gamma}(\hat{q}_p - u)}{1 - \hat{\gamma}}$ ein Schätzer von $CVaR_p(F)$ ist.

Der Beweis erfolgt mit Hilfe der folgenden 3 Schritte:

- (1) Sei $Z_1 \sim F$ eine ZV mit Verteilungsfunktion F und seien $u \geq u_0 > 0$ zwei Schwellwerte. Es gilt $e_{Z_1}(u) = e_{Z_2}(u - u_0)$ wobei $Z_2 \stackrel{d}{=} (Z_1 - u_0) | (Z_1 > u_0)$ und e_{Z_1} und e_{Z_2} die durchschnittlichen Exzessfunktionen von Z_1 bzw. Z_2 sind.

(2) Sei X eine ZV mit $X \sim GPD_{\gamma,0,\beta}$ und $0 < \gamma < 1$. Es gilt

$$CVaR_p(X) = q_p + \frac{\beta + \gamma q_p}{1 - \gamma},$$

wobei $q_p := VaR_p(X)$ das p -Quantil von X ist.

(3) Sei X eine ZV mit $X \sim F$. Die Randverteilung $\bar{F}(x)$ wird durch $\bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$ approximiert. Daraus folgt $F \approx \tilde{F}$ mit $\tilde{F} := 1 - \bar{F}(u)\bar{G}_{\gamma,0,\beta}(x-u)$. Für $q_p > u$ ist der CVaR der Approximation \tilde{F} folgendermaßen gegeben:

$$CVaR_p(\tilde{F}) = q_p + \frac{\beta + \gamma(q_p - u)}{1 - \gamma}$$