Die Verteilungen Φ_α, Ψ_α und Λ heißen *standard Extremwertverteilungen*. Verteilungen, die vom selben Typus wie Φ_α, Ψ_α oder Λ heißen *Extremwertverteilungen*.

Definition 16 *Die ZV X (oder die dazugehörige Verteilung) gehört zum maximalen Anziehungsgebiet der Extremwertverteilung H wenn die Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$ existieren, sodass*

$$\lim_{n\to\infty} a_n^{-1}(M_n - b_n) = H.$$

Notation: $X \in MDA(H)$ ($F \in MDA(H)$).

Theorem 8 *(Charakterisierung von MDA)*
$F \in MDA(H)$ mit normierenden Konstanten $a_n > 0$ und $b_n \in \mathbb{R}$, dann und nur dann, wenn

$$\lim_{n\to\infty} n\bar{F}(a_n x + b_n) = -\ln H(x), \forall x \in \mathbb{R}.$$

Für $H(x) = 0$ wird $-\ln H(x)$ durch ∞ ersetzt.

Hinweis zum Beweis: Satz 8 folgt vom Satz 5.

Es gibt auch Verteilungen die zu keinem maximalen Anziehungsgebiet einer Extremwertverteilung gehören!

Example 19 *(Die Poisson Verteilung)*
Sei $X \sim P(\lambda)$. D.h. $P(X = k) = e^{-\lambda}\lambda^k/k!$, $k \in \mathbb{N}_0$, $\lambda > 0$. Zeigen Sie, dass es keine Extremwertverteilung Z gibt für die $X \in MDA(Z)$.

Example 20 *(Maxima der Exponentialverteilung)*
Sei (X_k) eine Folge von i.i.d. ZV mit Verteilungsfunktion F, $F(x) = 1 - e^{-x}$ für $x \geq 0$. Zeigen Sie, dass $F \in MDA(\Lambda)$ mit normierenden Konstanten $a_n = 1$ und $b_n = \ln n$.

Example 21 *(Maxima der Cauchy-Verteilung)*
Sei (X_k) eine Folge von i.i.d. ZV mit Verteilungsfunktion F und Dichtefunktion f, $f(x) = (\pi(1 + x^2))^{-1}$ für $x \in \mathbb{R}$. Zeigen Sie, dass $F \in MDA(\Phi_1)$ mit normierenden Konstanten $a_n = n/\pi$ und $b_n = 0$.

Definition 17 *(Die Verallgemeinerte Extremwertverteilung)*
Die Verteilungsfunktion H_γ sei folgendermaßen gegeben:

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

wobei $1 + \gamma x > 0$. D.h. der Definitionsbereich von H_γ wird folgendermaßen gegeben:

$$\begin{array}{ll} x > -\gamma^{-1} & \text{wenn } \gamma > 0 \\ x < -\gamma^{-1} & \text{wenn } \gamma < 0 \\ x \in \mathbb{R} & \text{wenn } \gamma = 0 \end{array}$$

H_γ heißt verallgemeinerte standard Extremwertverteilung.

Theorem 9 *(Charakterisierung von $MDA(H_\gamma)$)*

- $F \in MDA(H_\gamma)$ *mit* $\gamma > 0 \iff F \in MDA(\Phi_\alpha)$ *mit* $\alpha = 1/\gamma > 0$.
- $F \in MDA(H_0) \iff F \in MDA(\Lambda)$.
- $F \in MDA(H_\gamma)$ *mit* $\gamma < 0 \iff F \in MDA(\Psi_\alpha)$ *mit* $\alpha = -1/\gamma > 0$.

Theorem 10 *($MDA(\Phi_\alpha)$, Gnedenko 1943)*

$F \in MDA(\Phi_\alpha)$ $(\alpha > 0) \iff \bar{F} \in RV_{-\alpha}$ $(\alpha > 0)$.
Wenn $F \in MDA(\Phi_\alpha)$, *dann* $\lim_{n\to\infty} a_n^{-1} M_n = \Phi_\alpha$ *mit* $a_n = F^{\leftarrow}(1-n^{-1})$.

Example 22 *Zeigen Sie, dass die untenstehenden Verteilungen dem $MDA(\Phi_\alpha)$ gehören und bestimmen Sie die normierenden Konstanten.*

- *Pareto:* $F(x) = 1 - x^{-\alpha}$, $x > 1$, $\alpha > 0$.
- *Cauchy:* $f(x) = (\pi(1 + x^2))^{-1}$, $x \in \mathbb{R}$.
- *Student:* $f(x) = \frac{\Gamma((\alpha+1)/2)}{\sqrt{\alpha\pi}\Gamma(\alpha/2)(1+x^2/\alpha)^{(\alpha+1)/2}}$, $\alpha \in \mathbb{N}$, $x \in \mathbb{R}$.
- *Loggamma:* $f(x) = \frac{\alpha^\beta}{\Gamma(\beta)}(\ln x)^{\beta-1} x^{-\alpha-1}$, $x > 1$, $\alpha, \beta > 0$.

Theorem 11 *($MDA(\Psi_\alpha)$, Gnedenko 1943)*

$F \in MDA(\Psi_\alpha)$ *($\alpha > 0$)* $\Longleftrightarrow$ $x_F := \sup\{x \in \mathbb{R}: F(x) < 1\} < \infty$ *und* $\bar{F}(x_F - x^{-1}) \in RV_{-\alpha}$ *($\alpha > 0$).*

Wenn $F \in MDA(\Psi_\alpha)$, *dann* $\lim_{n\to\infty} a_n^{-1}(M_n - x_F) = \Psi_\alpha$ *mit* $a_n = x_F - F^{\leftarrow}(1 - n^{-1})$.

Example 23 *Sei* $X \sim U(0,1)$. *Es gilt* $X \in MDA(\Psi_1)$ *mit* $a_n = 1/n$, $n \in \mathbb{N}$.

Theorem 12 *($MDA(\Lambda)$)*

Sei F *eine Verteilungsfunktion mit rechtem Endpunkt* $x_F \leq \infty$. $F \in MDA(\Lambda)$ *dann und nur dann wenn es ein* $z < x_F$ *existiert, sodass für* F *folgende Darstellung gilt:*

$$\bar{F}(x) = c(x) exp\Big\{ - \int_z^x \frac{g(t)}{a(t)} dt \Big\}, \forall x, z < x \leq x_F.$$

Für die Funktionen $c(x)$ *und* $g(x)$ *gilt* $\lim_{x\uparrow x_F} c(x) = c > 0$ *und* $\lim_{t\uparrow x_F} g(t) = 1$, *und* $a(t)$ *ist eine positive absolut stetige Funktion, sodass* $\lim_{t\uparrow x_F} a'(t) = 0$.

Theorem 13 *($MDA(\Lambda)$, alternative Charakterisierung)*

Eine Verteilungsfunktion F gehört zu $MDA(\Lambda)$ dann und nur dann wenn es eine positive Funktion $\tilde{a}$ existiert, sodass

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x + u\tilde{a}(x))}{\bar{F}(x)} = e^{-u}, \forall u \in \mathbb{R}$$

Eine mögliche Wahl für $\tilde{a}$ ist $\tilde{a}(x) = a(x)$

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt$$

Die Funktion $a(x)$ heißt durchschnittliche Überschußfunktion (mean excess function):

$$a(x) = E(X - x | X > x), \forall x \leq x_F$$

Einige Verteilungen, die dem $MDA(\Lambda)$ gehören:

- Normal: $F(x) = (2\pi)^{-1/2} \exp\{-x^2/2\}$, $x \in \mathbb{R}$.
- Exponential: $f(x) = \lambda^{-1} \exp\{-\lambda x\}$, $x > 0$, $\lambda > 0$.
- Lognormal: $f(x) = (2\pi x^2)^{-1/2} \exp\{-(\ln x)^2/2\}$, $x > 0$.
- Gamma: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}$, $x > 0$, $\alpha, \beta > 0$.

Graphische Methoden zur Untersuchung des Verteilungsrandes

- Histogramm

- Quantil-Quantil Plots

  $X_1, X_2, \ldots, X_n$ sind i.i.d. ZV mit einer unbekannten Verteilung $\tilde{F}$. Es wird vermutet, dass $\tilde{F}$ am Rand von einer bekannten Verteilung F approximiert wird. Wie kann man diese Vermutung testen?

  Sei $X_{n,n} \le X_{n-1,n} \le \ldots \le X_{1,n}$ eine sortierte Stichprobe aus X_1, X_2,..., X_n.

  qq-plot: $\{(X_{k,n}, F^{\leftarrow}(\frac{n-k+1}{n+1}))\colon k = 1, 2, \ldots, n\}$.

  Bei einer plausiblen Vermutung stellt der qq-plot eine einigermaßen lineare Abhängigkeit dar. Diese Eigenschaft bleibt auch dann erhalten wenn die echte Verteilung und die Referenz-Verteilung nicht übereinstimmen, sondern vom selben Typus sind.

  Faustregel: Je größer das Quantil um so mehr “heavy tailed” ist die Verteilung!

Der Hill Schätzer

Seien $X_1, X_2, \ldots, X_n$ i.i.d. mit Verteilungsfunktion F, sodass $\bar{F} \in RV_{-\alpha}$, $\alpha > 0$, d.h. $\bar{F}(x) = x^{-\alpha} L(x)$ mit $L \in RV_0$.

Ziel: Schätzung von α!

Theorem 14 *(Satz von Karamata)*
Sei L eine langsam variierende und lokal beschränkte Funktion auf $[x_0, +\infty)$ für ein $x_0 \in \mathbb{R}$. Dann gilt:

(a) Für $\kappa > -1$:

$$\int_{x_o}^{x} t^{\kappa} L(t) dt \sim K(x_0) + \frac{1}{\kappa + 1} x^{\kappa+1} L(x) \text{ für } x \to \infty$$

($K(x_0)$ *ist eine von x_0 abhängige Konstante.*)

(b) Für $\kappa < -1$:

$$\int_{x}^{+\infty} t^{\kappa} L(t) dt \sim -\frac{1}{\kappa + 1} x^{\kappa+1} L(x) \text{ für } x \to \infty$$

Beweis in Bingham et al. 1987.

Annahme: L ist lokal beschränkt in $[u, +\infty)$.

Aus dem Satz von Karamata folgt:

$$E\left(\ln(X) - \ln(u) | \ln(X) > \ln(u)\right) = \lim_{u\to\infty} \frac{1}{\bar{F}(u)} \int_u^\infty (\ln x - \ln u) dF(x) = \alpha^{-1}. \tag{8}$$

Für die empirische Verteilung $F_n(x) = \frac{1}{n}\sum_{k=1}^n I_{[x_k,\infty)}(x)$ und eine hohe stichprobenabhängige Schwelle $x_{k,n}$, erhält man:

$$E\left(\ln(X) - \ln(x_{k,n}) | \ln(X) > \ln(x_{k,n})\right) \approx$$

$$\frac{1}{\bar{F}_n(x_{k,n})} \int_{X_{k,n}}^\infty (\ln x - \ln x_{k,n}) dF_n(x) = \frac{1}{k-1} \sum_{j=1}^{k-1} (\ln x_{j,n} - \ln x_{k,n}).$$

Wenn $k = k(n) \to \infty$ und $k/n \to 0$, dann $x_{k,n} \to \infty$ für $n \to \infty$, und dann folgt aus (8):