KAPITEL 3

Satz von Fisher–Tippett–Gnedenko

In diesem Kapitel beweisen wir den Satz von Fisher–Tippett–Gnedenko, der die Extremwertverteilungen beschreibt.

Satz 3.0.1 (Satz von Fisher–Tippett (1928), Gnedenko (1943))**.** Eine Verteilungsfunktion G ist eine Extremwertverteilung genau dann, wenn einer der drei folgenden Fälle eintritt:

(1) G ist vom Gumbel–Typ, d.h. $G(t) = \Lambda\left(\frac{t-\mu}{\sigma}\right)$ für $\mu \in \mathbb{R}$, $\sigma > 0$.
(2) G ist vom Fréchet–Typ, d.h. $G(t) = \Phi_\alpha\left(\frac{t-\mu}{\sigma}\right)$ für $\alpha > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.
(3) G ist vom Weibull–Typ, d.h. $G(t) = \Psi_\alpha\left(\frac{t-\mu}{\sigma}\right)$ für $\alpha > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.

Für den Beweis benötigen wir einige Hilfsmittel.

3.1. Eindeutigkeit der Normierungskonstanten

Seien $Z_1, Z_2, \ldots$ Zufallsvariablen und $a_n > 0$ und $b_n \in \mathbb{R}$ Folgen von Normierungskonstanten mit

$$\frac{Z_n - b_n}{a_n} \xrightarrow[n\to\infty]{d} Z. \tag{3.1.1}$$

Wir stellen uns die Frage, wie stark wir die Konstanten a_n, b_n verändern können, ohne dass die Konvergenz in (3.1.1) zerstört wird.

Proposition 3.1.1. Es gelte (3.1.1). Seien $\tilde{a}_n > 0$ und $\tilde{b}_n \in \mathbb{R}$ zwei weitere Folgen mit

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{a_n} = 1, \quad \lim_{n\to\infty} \frac{\tilde{b}_n - b_n}{a_n} = 0.$$

Dann gilt auch

$$\frac{Z_n - \tilde{b}_n}{\tilde{a}_n} \xrightarrow[n\to\infty]{d} Z.$$

Beweis. Es gilt

$$\frac{Z_n - \tilde{b}_n}{\tilde{a}_n} = \frac{a_n}{\tilde{a}_n} \cdot \frac{Z_n - \tilde{b}_n}{a_n} = \frac{a_n}{\tilde{a}_n} \cdot \left(\frac{Z_n - b_n}{a_n} - \frac{\tilde{b}_n - b_n}{a_n}\right). \tag{3.1.2}$$

Nach dem Lemma von Slutsky konvergiert die rechte Seite gegen Z in Verteilung. □

Nun stellen wir eine allgemeinere Frage: Wie stark können wir die Folgen a_n, b_n verändern, so dass die Konvergenz in (3.1.1) bestehen bleibt, allerdings eventuell mit einem anderen Grenzwert als Z?

Proposition 3.1.2. Es gelte (3.1.1). Seien $\tilde{a}_n > 0$ und $\tilde{b}_n \in \mathbb{R}$ zwei weitere Folgen mit

$$\lim_{n\to\infty} \frac{\tilde{a}_n}{a_n} = a \in (0,\infty), \quad \lim_{n\to\infty} \frac{\tilde{b}_n - b_n}{a_n} = b \in \mathbb{R}.$$

Dann gilt auch

$$\frac{Z_n - \tilde{b}_n}{\tilde{a}_n} \xrightarrow[n\to\infty]{d} \frac{Z-b}{a}.$$

Beweis. Wir können nach wie vor (3.1.2) verwenden. Nach dem Lemma von Slutsky konvergiert die rechte Seite gegen $(Z-b)/a$ in Verteilung. □

Satz 3.1.3 ("Convergence of types theorem", Chintschin)**.** Seien $Z_1, Z_2, \ldots$ Zufallsvariablen und $a_n > 0$, $b_n \in \mathbb{R}$ sowie $\tilde{a}_n > 0$, $\tilde{b}_n \in \mathbb{R}$ Normierungsfolgen mit

$$\frac{Z_n - b_n}{a_n} \xrightarrow[n\to\infty]{d} Z, \quad \frac{Z_n - b_n}{a_n} \xrightarrow[n\to\infty]{d} \tilde{Z},$$

wobei die Zufallsvariablen $Z, \tilde{Z}$ nicht degeneriert seien. Dann existieren die Grenzwerte

$$a := \lim_{n\to\infty} \frac{\tilde{a}_n}{a_n} \in (0,\infty), \quad b := \lim_{n\to\infty} \frac{\tilde{b}_n - b_n}{a_n} \in \mathbb{R} \tag{3.1.3}$$

und es gilt $\tilde{Z} \stackrel{d}{=} (Z-b)/a$.

Beweis. Weggelassen. Referenz: P. Billingsley, *Probability and measure*, Seite 193, Thm. 4.2. □

Bemerkung 3.1.4. Bezeichnen wir mit $G_1, G_2, \ldots$ bzw. $G, \tilde{G}$ die Verteilungsfunktionen von $Z_1, Z_2, \ldots$ bzw. $Z, \tilde{Z}$, so können wir den obigen Satz auch wie folgt formulieren: Aus

$$\lim_{n\to\infty} G_n(a_n t + b_n) = G(t), \quad \lim_{n\to\infty} G_n(\tilde{a}_n t + \tilde{b}_n) = \tilde{G}(t),$$

mit nicht degenerierten Grenzwertverteilungen G und $\tilde{G}$ folgt, dass (3.1.3) gilt und dass $\tilde{G}(t) = G(at+b)$.

Lemma 3.1.5. Sei F eine nicht degenerierte Verteilungsfunktion und $c > 0$, $d \in \mathbb{R}$ Konstanten mit $F(ct+d) = F(t)$ für alle $t \in \mathbb{R}$. Dann ist $c = 1$, $d = 0$.

Beweis. Weggelassen. Referenz: P. Billingsley, ab Seite 193. □

3.2. Max-stabile Verteilungen

Definition 3.2.1. Eine nicht degenerierte Verteilungsfunktion G heißt **max-stabil**, falls es für alle $n \in \mathbb{N}$ Konstanten $c_n > 0$ und $d_n \in \mathbb{R}$ gibt mit

$$G^n(c_n t + d_n) = G(t).$$

Mit anderen Worten, für jedes $n \in \mathbb{N}$ ist G^n vom gleichen Typ wie G.

Sind $X_1, X_2, \ldots$ u.i.v. Zufallsvariablen mit einer max-stabilen Verteilung, dann gilt für alle $n \in \mathbb{N}$

$$\frac{\max\{X_1, \ldots, X_n\} - d_n}{c_n} \stackrel{d}{=} X_1.$$

Das heißt, das Maximum von n u.i.v. Zufallsvariablen mit einer max-stabilen Verteilung hat bis auf eine affine Transformation die gleiche Verteilung wie eine einzige Zufallsvariable.

Beispiel 3.2.2. Die Gumbel–Verteilungsfunktion $\Lambda(t) = \mathrm{e}^{-\mathrm{e}^{-t}}$ ist max-stabil, denn

$$\Lambda^n(t + \log n) = \mathrm{e}^{-n\mathrm{e}^{-(t+\log n)}} = \mathrm{e}^{-\mathrm{e}^{-t}} = \Lambda(t).$$

Analog lässt sich zeigen, dass Fréchet–Verteilung Φ_α und Weibull–Verteilung Ψ_α max-stabil sind.

Die Klasse der max-stabilen Verteilungen stimmt mit der Klasse der Extremwertverteilungen überein:

Satz 3.2.3. Eine Verteilungsfunktion G ist max-stabil genau dann, wenn G eine Extremwertverteilung ist.

Beweis. Schritt 1. "$\Rightarrow$" Sei G max-stabil. Dann gibt es $c_n > 0, d_n \in \mathbb{R}$, so dass $G^n(c_n t + d_n) = G(t)$. Es gilt also für alle $t \in \mathbb{R}$

$$\lim_{n\to\infty} G^n(c_n t + d_n) = G(t),$$

weshalb G eine Extremwertverteilung ist. Wir haben übrigens gezeigt, dass die Verteilungsfunktion G in ihrem eigenen Max–Anziehungsbereich liegt.

Schritt 2. "$\Leftarrow$" Sei G eine Extremwertverteilung. Dann gibt es eine Verteilungsfunktion F und $a_n > 0, b_n \in \mathbb{R}$, so dass

$$\lim_{n\to\infty} F^n(a_n t + b_n) = G(t) \tag{3.2.1}$$

für alle Stetigkeitspunkte von G. Damit gilt für alle $k \in \mathbb{N}$

$$\lim_{n\to\infty} F^{nk}(a_{nk} t + b_{nk}) = G(t).$$

Indem wir die k-te Wurzel ziehen, erhalten wir

$$\lim_{n\to\infty} F^n(a_{nk} t + b_{nk}) = G^{1/k}(t). \tag{3.2.2}$$

Wir wenden nun den Chintschin–Satz 3.1.3 auf (3.2.1) und (3.2.2) an. Es folgt, dass G und $G^{1/k}$ vom gleichen Typ sind, d.h. es gibt $c_k > 0$ und $d_k \in \mathbb{R}$ mit

$$G^{1/k}(t) = G(c_k t + d_k) \quad \text{bzw.} \quad G(t) = G^k(c_k t + d_k).$$

Das bedeutet aber, dass G max-stabil ist. $\square$

Von nun an besteht unser Ziel darin, die max-stabilen Verteilungen zu beschreiben. Die nächste Proposition zeigt, dass die Eigenschaft $G^n(c_n t + d_n) = G(t)$ sogar auf nicht-ganzzahlige Werte von n erweitert werden kann.

Proposition 3.2.4. Sei G eine max-stabile Verteilungsfunktion. Dann gibt es messbare Funktionen $c : (0, \infty) \to (0, \infty)$ und $d : (0, \infty) \to \mathbb{R}$, so dass für alle $s > 0$ (nicht notwendigerweise ganzzahlig) gilt:

$$G^s(c(s)t + d(s)) = G(t). \tag{3.2.3}$$

Beweis. Wir bezeichnen mit $[t]$ die Gaußklammer einer reellen Zahl t:

$$[t] = \max\{n \in \mathbb{Z} | n \le t\}.$$

Sei G eine max-stabile Verteilungsfunktion. Dann gibt es $c_n > 0, d_n \in \mathbb{R}$, so dass für alle $n \in \mathbb{N}$

$$G^n(c_n t + d_n) = G(t). \tag{3.2.4}$$

Für beliebiges $s > 0$ folgt daraus, dass

$$G^{[ns]}(c_{[ns]} t + d_{[ns]}) = G(t).$$

Daraus ergibt sich, dass für alle $t \in \mathbb{R}$

$$G^n(c_{[ns]} t + d_{[ns]}) = (G^{[ns]}(c_{[ns]} t + d_{[ns]}))^{\frac{n}{[ns]}} = G^{\frac{n}{[ns]}}(t) \underset{n\to\infty}{\longrightarrow} G^{1/s}(t). \tag{3.2.5}$$

Gleichzeitig gilt aber wegen (3.2.4) auch

$$G^n(c_n t + d_n) \underset{n\to\infty}{\longrightarrow} G(t). \tag{3.2.6}$$

Mit Satz 3.1.3 folgt aus (3.2.5) und (3.2.6), dass die folgenden Grenzwerte existieren:

$$c(s) := \lim_{n\to\infty} \frac{c_{[ns]}}{c_n} \in (0, \infty), \quad d(s) := \lim_{n\to\infty} \frac{d_{[ns]} - d_n}{c_n} \in \mathbb{R}, \tag{3.2.7}$$

und dass $G^{1/s}(t) = G(c(s)t + d(s))$. Insgesamt folgt also $G(t) = G^s(c(s)t + d(s))$.
Außerdem folgt aus der Darstellung (3.2.7), dass die Funktionen c und d als punktweise Grenzwerte von Folgen messbarer Funktionen, selber messbar sind. $\square$

3.3. Charakterisierung der max-stabilen Verteilungen

Wegen Satz 3.2.3 können wir den Satz von Fisher–Tippett–Gnedenko nun wie folgt formulieren:

Satz 3.3.1. Jede max-stabile Verteilungsfunktion G ist vom gleichen Typ wie eine der folgenden Verteilungen: Gumbel Λ, Fréchet Φ_α mit $\alpha > 0$ oder Weibull Ψ_α mit $\alpha > 0$.

Beweis. Sei G eine max-stabile Verteilungsfunktion. Laut Proposition 3.2.4 gibt es messbare Funktionen $c(s) > 0$, $d(s) \in \mathbb{R}$ mit

$$G^s(c(s)t + d(s)) = G(t) \text{ für alle } s > 0,\ t \in \mathbb{R}. \tag{3.3.1}$$

Im Folgenden werden wir diese Funktionalgleichung lösen. Zuerst werden wir G eliminieren und die Funktionen c und d bestimmen.

Schritt 1. Für $s > 0$ betrachte die affine Transformation $\varphi_s : \mathbb{R} \to \mathbb{R}$ mit

$$\varphi_s(t) = c(s)t + d(s), \quad t \in \mathbb{R}.$$

Es gilt mit dieser Notation für alle $s > 0$ und $t \in \mathbb{R}$, dass

$$G(\varphi_s(t)) = G^{1/s}(t). \tag{3.3.2}$$

Wir zeigen, dass die Abbildung $s \mapsto \varphi_s$ ein Homomorphismus aus der multiplikativen Gruppe der positiven reellen Zahlen $(\mathbb{R}_{>0}, \cdot)$ in die Gruppe der affinen Transformationen von $\mathbb{R}$ ist, d.h. für alle $s_1, s_2 > 0$, $t \in \mathbb{R}$

$$\varphi_{s_1 s_2}(t) = \varphi_{s_2}(\varphi_{s_1}(t)). \tag{3.3.3}$$

Indem wir in (3.3.2) mehrmals benutzen, erhalten wir, dass

$$G(\varphi_{s_1 s_2}(t)) = G^{1/(s_1 s_2)}(t) = (G^{1/s_1}(t))^{1/s_2} = (G(\varphi_{s_1}(t)))^{1/s_2} = G(\varphi_{s_2}(\varphi_{s_1}(t))).$$

Aus Lemma 3.1.5 folgt (3.3.3).

Schritt 2. Mit $\varphi_s(t) = c(s)t + d(s)$ erhalten wir

$$\begin{aligned} \varphi_{s_1 s_2}(t) &= c(s_1 s_2)t + d(s_1 s_2), \\ \varphi_{s_2}(\varphi_{s_1}(t)) &= c(s_1)c(s_2)t + c(s_2)d(s_1) + d(s_2). \end{aligned}$$

Somit führt (3.3.3) zum folgenden System von Funktionalgleichungen: Für alle $s_1, s_2 > 0$ und alle $t \in \mathbb{R}$,

$$\begin{cases} c(s_1 s_2) = c(s_1)c(s_2), \\ d(s_1 s_2) = c(s_2)d(s_1) + d(s_2). \end{cases}$$

Die erste Gleichung ist eine Cauchy–Funktionalgleichung in multiplikativer Form. Da c eine messbare Funktion ist, gibt es ein $\rho \in \mathbb{R}$ mit

$$c(s) = s^\rho, \quad s > 0.$$

Schritt 3. Wir betrachten zunächst den Fall $\rho = 0$, also $c(s) = 1$. Die Funktionalgleichung für d sieht folgendermaßen aus: Für alle $s_1, s_2 > 0$

$$d(s_1 s_2) = d(s_1) + d(s_2).$$

Betrachte die Funktion $h(x) = d(\mathrm{e}^x)$, $x \in \mathbb{R}$. Für diese gilt

$$h(x_1 + x_2) = d(\mathrm{e}^{x_1 + x_2}) = d(\mathrm{e}^{x_1}\mathrm{e}^{x_2}) = d(\mathrm{e}^{x_1}) + d(\mathrm{e}^{x_2}) = h(x_1) + h(x_2).$$

Somit ist h additiv. Da h messbar ist, folgt aus dem Satz von Ostrowski, dass $h(x) = ax$ für ein $a \in \mathbb{R}$ und somit $d(s) = a \log s$. Im Fall $\rho = 0$ haben wir also die folgende Lösung erhalten:

$$c(s) = 1, \quad d(s) = a \log s.$$

Die Gleichung (3.3.1) für die Verteilungsfunktion G vereinfacht sich also zu

$$G(t + a \log s) = G^{1/s}(t).$$

Der Fall $a = 0$ ist ausgeschlossen. In diesem Fall wäre nämlich $G(t) = G^{1/s}(t)$ für alle $s > 0$, was bedeuten würde, dass $G(t)$ nur die Werte 0 oder 1 annimmt. Widerspruch, denn G ist nicht degeneriert. Sei also $a \neq 0$. Mit $t = 0$ und $y = a \log s \in \mathbb{R}$ erhalten wir

$$G(y) = G^{1/s}(0) = (G(0))^{\mathrm{e}^{-y/a}} = \exp\left\{(\log G(0)) \cdot \mathrm{e}^{-\frac{y}{a}}\right\} = \exp\left\{-\mathrm{e}^{-\frac{y}{a} + \log(-\log G(0))}\right\}.$$

Das gilt wegen $a \neq 0$ für alle $y \in \mathbb{R}$. Es sei bemerkt, dass $\log G(0) < 0$ und dass die Fälle $G(0) = 1$ oder 0 ausgeschlossen sind, denn sonst wäre G identisch 0 oder 1, Widerspruch. Somit ist G vom Gumbel–Typ.

Schritt 4. Nun betrachten wir den Fall $\rho \neq 0$. Es gilt $c(s) = s^\rho$ und die Gleichung für d nimmt die folgende Form an: Für alle $s_1, s_2 > 0$

$$d(s_1 s_2) = s_2^\rho d(s_1) + d(s_2).$$

Indem wir s_1 und s_2 vertauschen, erhalten wir

$$d(s_2 s_1) = s_1^\rho d(s_2) + d(s_1).$$

Somit ergibt sich für alle $s_1, s_2 > 0$

$$s_2^\rho d(s_1) + d(s_2) = s_1^\rho d(s_2) + d(s_1).$$

Sei nun $s_2 = 2$. Aus $\rho \neq 0$ folgt, dass $2^\rho - 1 \neq 0$ und wir erhalten, dass für alle $s_1 > 0$

$$d(s_1) = \frac{d(2)}{2^\rho - 1}(s_1^\rho - 1) = (s_1^\rho - 1)\mu,$$

wobei $\mu = \frac{d(2)}{2^\rho - 1} \in \mathbb{R}$. Wir haben die Funktionen c und d bestimmt:

$$c(s) = s^\rho, \quad d(s) = (s^\rho - 1)\mu,$$

wobei $\rho \in \mathbb{R}$ und $\mu \in \mathbb{R}$ zwei Parameter sind.
Die Gleichung (3.3.1) für die Verteilungsfunktion G nimmt somit die folgende Gestalt an: Für alle $s > 0$, $t \in \mathbb{R}$,

$$G^s(s^\rho t + (s^\rho - 1)\mu) = G(t).$$

Betrachte die Verteilungsfunktion $H(u) = G(u - \mu)$, $u \in \mathbb{R}$. Die Gleichung für H sieht wie folgt aus: Für alle $s > 0$, $t \in \mathbb{R}$,

$$H^s(s^\rho t) = H(t).$$

Mit $t = 0$ erhalten wir $H^s(0) = H(0)$ für alle $s > 0$, somit ist $H(0) = 0$ oder $H(0) = 1$.

Fall 1. Sei $H(0) = 0$. Da H eine Verteilungsfunktion ist, gilt $H(y) = 0$ für alle $y \leq 0$. Sei nun $y > 0$. Mit $t = 1$ und $s = y^{1/\rho} > 0$ erhalten wir

$$H(y) = H^{1/s}(1) = H^{y^{-1/\rho}}(1) = \exp\left\{(\log H(1)) \cdot y^{-1/\rho}\right\}, \quad y > 0.$$

Es sei bemerkt, dass $\log H(1) < 0$. Es folgt, dass H und somit auch G vom gleichen Typ wie die Fréchet–Verteilung $\Phi_{1/\rho}$ ist.

Fall 2. Sei $H(0) = 1$. Da H eine Verteilungsfunktion ist, gilt $H(y) = 1$ für alle $y \geq 0$. Sei nun $y < 0$. Mit $t = -1$ und $s = (-y)^{1/\rho} > 0$ erhalten wir

$$H(y) = H^{1/s}(-1) = H^{(-y)^{-1/\rho}}(-1) = \exp\left\{(\log H(-1)) \cdot (-y)^{-1/\rho}\right\}, \quad y < 0.$$

Es sei bemerkt, dass $\log H(-1) < 0$. Es folgt, dass H und somit auch G vom gleichen Typ wie die Weibull–Verteilung $\Psi_{-1/\rho}$ ist. $\square$