

1 Extremwertverteilungen

Seien X_1, X_2, \dots unabhängige und identisch verteilte (u.i.v.) Zufallsvariablen. Mit M_n bezeichnen wir das **Maximum** von X_1, \dots, X_n :

$$M_n = \max\{X_1, \dots, X_n\}, \quad n \in \mathbb{N}.$$

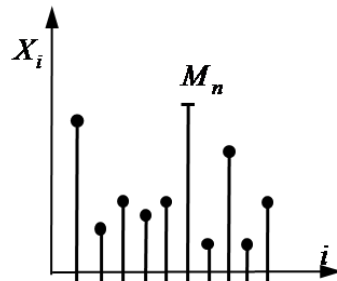


Abbildung 1.1: Veranschaulichung M_n

Die **Verteilungsfunktion** F und die **Tailfunktion** \bar{F} von X_i sind definiert durch

$$F(t) = \mathbb{P}[X_i \leq t], \quad \bar{F}(t) = 1 - F(t) = \mathbb{P}[X_i > t], \quad t \in \mathbb{R}.$$

Die Verteilungsfunktion F ist monoton steigend. Die Tailfunktion \bar{F} hingegen ist monoton fallend.

Theorem 1.1. Die Verteilungsfunktion von M_n ist gegeben durch

$$\mathbb{P}[M_n \leq t] = F^n(t) \quad \text{für alle } t \in \mathbb{R}.$$

Beweis. Aus der Definition des Maximums M_n folgt

$$\mathbb{P}[M_n \leq t] = \mathbb{P}[\max\{X_1, \dots, X_n\} \leq t] = \mathbb{P}[X_1 \leq t, \dots, X_n \leq t].$$

Da X_1, \dots, X_n unabhängig und identisch verteilt sind, gilt

$$\mathbb{P}[X_1 \leq t, \dots, X_n \leq t] = \mathbb{P}[X_1 \leq t] \cdot \dots \cdot \mathbb{P}[X_n \leq t] = F^n(t).$$

Zusammen ergibt sich $\mathbb{P}[M_n \leq t] = F^n(t)$. □

Wir werden uns für die Eigenschaften von M_n für große Werte von n interessieren. Im folgenden Satz berechnen wir den Wert, gegen den die Zufallsvariable M_n für $n \rightarrow \infty$ in Wahrscheinlichkeit konvergiert. Wir definieren den rechten Endpunkt der Verteilungsfunktion F durch

$$x^* = \sup\{t \in \mathbb{R} | F(t) < 1\}.$$

Es kann passieren, dass x^* den Wert $+\infty$ annimmt.

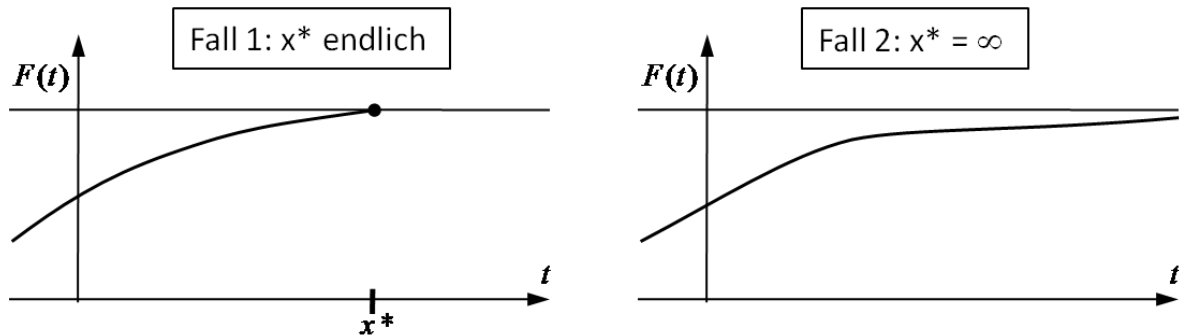


Abbildung 1.2: Veranschaulichung x^*

Theorem 1.2. Für $n \rightarrow \infty$ konvergiert die Zufallsvariable M_n gegen den Wert x^* in Wahrscheinlichkeit.

Beweis. Sei zuerst x^* endlich. Für jedes $\varepsilon > 0$ gilt $F(x^* - \varepsilon) < 1$, wobei die Ungleichung strikt ist. Aus Satz 1.1 folgt, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n \leq x^* - \varepsilon] = \lim_{n \rightarrow \infty} F^n(x^* - \varepsilon) = 0.$$

Außerdem gilt $\mathbb{P}[M_n \geq x^* + \varepsilon] = 0$. Es folgt

$$\lim_{n \rightarrow \infty} \mathbb{P}[|M_n - x^*| \geq \varepsilon] = \lim_{n \rightarrow \infty} (\mathbb{P}[M_n \leq x^* - \varepsilon] + \mathbb{P}[M_n \geq x^* + \varepsilon]) = 0$$

Somit gilt $M_n \xrightarrow{P} x^*$.

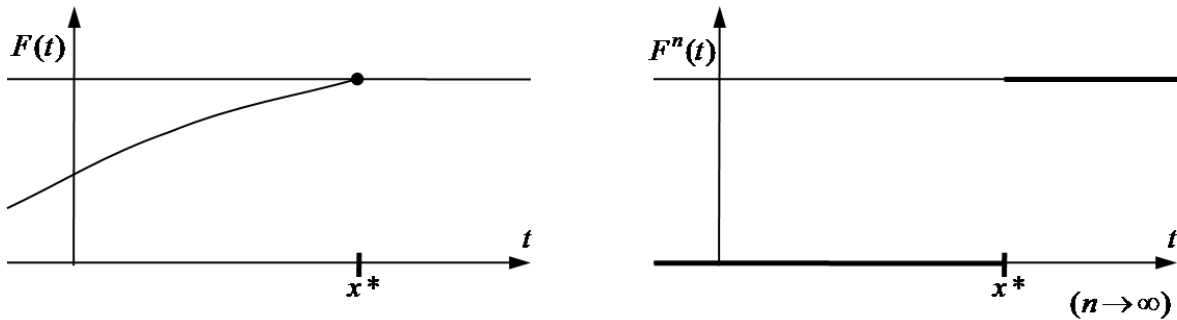
Sei nun $x^* = +\infty$. Wir zeigen, dass für jedes $A \in \mathbb{R}$, das beliebig groß sein kann, $\lim_{n \rightarrow \infty} \mathbb{P}[M_n < A] = 0$ gilt. Aus $x^* = \infty$ folgt, dass $F(A) < 1$, wobei die Ungleichung strikt ist. Mit Satz 1.1 folgt

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n < A] = \lim_{n \rightarrow \infty} F^n(A) = 0.$$

Somit gilt $M_n \xrightarrow{P} +\infty$. □

Aufgabe 1.3. Zeigen Sie, dass M_n sogar fast sicher gegen x^* konvergiert.

Satz 1.2 gibt nur eine sehr grobe Information darüber, wie die Verteilungsfunktion von M_n für $n \rightarrow \infty$ aussieht:



Zum Vergleich besagt das schwache Gesetz der großen Zahlen, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit endlichem Erwartungswert gilt

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow{P} \mathbb{E}X_1.$$

Dies gibt nur eine sehr grobe Vorstellung darüber, wie die Verteilungsfunktion von $X_1 + \dots + X_n$ aussieht. Eine präzise Aussage über die Verteilungsfunktion von $X_1 + \dots + X_n$ liefert der zentrale Grenzwertsatz. Er besagt, dass für u.i.v. Zufallsvariablen X_1, X_2, \dots mit endlichem Erwartungswert μ und endlicher Varianz $\sigma^2 > 0$

$$\frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1)$$

im Sinne der Verteilungskonvergenz, wobei $N(0, 1)$ eine standardnormale Verteilung bezeichnet. Wir stellen nun eine ähnliche Frage für das Maximum M_n .

Frage 1.4. Seien X_1, X_2, \dots u.i.v. Zufallsvariablen mit Verteilungsfunktion F . Gibt es Folgen von Konstanten $a_n \neq 0$, $b_n \in \mathbb{R}$ und eine Verteilungsfunktion G , sodass für $n \rightarrow \infty$ die folgende Verteilungskonvergenz gilt:

$$\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \xrightarrow{d} G? \tag{1.1}$$

Die Verteilungskonvergenz in (1.1) bedeutet, dass für alle $t \in \mathbb{R}$ in denen G stetig ist, gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \leq t \right] = G(t).$$

Definition 1.5. Eine Verteilungsfunktion G , für die (1.1) bei irgendeiner Wahl der Normierungskonstanten a_n, b_n und der Verteilungsfunktion F gilt, heißt **Extremwertverteilung**. Falls (1.1) gilt, sagen wir, dass F im **Max-Anziehungsbereich** von G liegt.

Bemerkung 1.6. Durch die “falsche” Wahl der Konstanten a_n, b_n kann man immer erreichen, dass (1.1) mit einer “ausgearteten” Verteilungsfunktion G der Form

$$G(t) = \begin{cases} 0, & t < t_0 \\ 1, & t \geq t_0 \end{cases} \quad (1.2)$$

gilt. Eine gemäß G verteilte Zufallsvariable nimmt nur den Wert t_0 mit Wahrscheinlichkeit 1 an. Zum Beispiel kann man a_n so groß wählen, dass M_n/a_n gegen 0 in Verteilung konvergiert. Deshalb werden wir den Fall einer Verteilungsfunktion G der Form (1.2) immer ausschließen. Eine ausgeartete Verteilung gilt nicht als Extremwertverteilung.

Mit der obigen Definition ist es nicht klar, ob Extremwertverteilungen überhaupt existieren. Im Folgenden betrachten wir 3 Beispiele, die 3 verschiedene Extremwertverteilungen (oder sogar Familien von Extremwertverteilungen) liefern.

Theorem 1.7. Seien X_i exponentialverteilt mit Parameter 1, d.h. $F(t) = 1 - e^{-t}$ für $t \geq 0$. Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}[M_n - \log n \leq t] = e^{-e^{-t}} =: G(t), \quad t \in \mathbb{R}.$$

Mit anderen Worten konvergiert die Zufallsvariable $M_n - \log n$ in Verteilung gegen G .

Beweis. Sei $t \in \mathbb{R}$ beliebig. Mit Satz 1.1 gilt

$$\mathbb{P}[M_n - \log n \leq t] = \mathbb{P}[M_n \leq t + \log n] = F^n(t + \log n).$$

Die Zufallsvariablen X_i sind exponentialverteilt und $t + \log n > 0$ für n groß genug. Es folgt

$$\mathbb{P}[M_n - \log n \leq t] = (1 - e^{-(t+\log n)})^n = \left(1 - \frac{e^{-t}}{n}\right)^n \xrightarrow[n \rightarrow \infty]{} e^{-e^{-t}}.$$

Somit gilt $M_n - \log n \xrightarrow{d} G$. □

Definition 1.8. Die Verteilungsfunktion $G(t) = e^{-e^{-t}}$ heißt **Gumbel-Verteilung**.

Bemerkung 1.9. Gemäß Satz 1.7 liegt die Exponentialverteilung im Max-Anziehungsbereich der Gumbel-Verteilung. Man kann Satz 1.7 wie folgt interpretieren: für großes n nimmt das Maximum M_n Werte an, die sich von dem Wert $\log n$ um eine approximativ Gumbel-verteilte “Fluktuation” unterscheiden.

Theorem 1.10. Seien X_i Pareto-verteilt mit Parameter $\alpha > 0$, d.h.

$$F(t) = \begin{cases} 1 - \frac{1}{t^\alpha}, & t \geq 1 \\ 0, & t \leq 1. \end{cases}$$

Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\frac{M_n}{n^{1/\alpha}} \leq t \right] = \begin{cases} e^{-\frac{1}{t^\alpha}}, & t \geq 0 \\ 0, & t \leq 0 \end{cases} =: \Phi_\alpha(t). \quad (1.3)$$

Mit anderen Worten konvergiert die Zufallsvariable $M_n/n^{1/\alpha}$ in Verteilung gegen Φ_α .

Beweis. Sei $t > 0$ beliebig. Mit Satz 1.1 erhalten wir

$$\mathbb{P} \left[\frac{M_n}{n^{1/\alpha}} \leq t \right] = \mathbb{P}[M_n \leq tn^{1/\alpha}] = F^n(tn^{1/\alpha}).$$

Da die Zufallsvariablen X_i Pareto-verteilt sind, folgt

$$\mathbb{P} \left[\frac{M_n}{n^{1/\alpha}} \leq t \right] = \left(1 - \frac{1}{(tn^{1/\alpha})^\alpha} \right)^n = \left(1 - \frac{1}{t^\alpha n} \right)^n \xrightarrow{n \rightarrow \infty} e^{-\frac{1}{t^\alpha}}.$$

Für $t < 0$ gilt $\mathbb{P}[M_n/n^{1/\alpha} \leq t] = 0$. Die Behauptung folgt. \square

Definition 1.11. Die Verteilungsfunktion Φ_α in (1.3) heißt **Fréchet-Verteilung** mit Parameter $\alpha > 0$.

Bemerkung 1.12. Man kann Satz 1.10 wie folgt interpretieren: für großes n nimmt das Maximum M_n sehr große Werte auf der Skala $n^{1/\alpha}$ an. Reskaliert man M_n mit dem Faktor $n^{-1/\alpha}$, so erhält man approximativ Fréchet-verteilte Werte.

Theorem 1.13. Die Verteilungsfunktion F von X_i habe die Form

$$F(t) = \begin{cases} 0, & t \leq 0, \\ 1 - (1 - t)^\alpha, & t \in [0, 1], \\ 1, & t \geq 1, \end{cases}$$

wobei $\alpha > 0$ ein Parameter ist. Dann gilt

$$\lim_{n \rightarrow \infty} \mathbb{P}[n^{1/\alpha}(M_n - 1) \leq t] = \begin{cases} e^{-(-t)^\alpha}, & t \leq 0 \\ 1, & t \geq 0 \end{cases} =: \Psi_\alpha(t). \quad (1.4)$$

Mit anderen Worten konvergiert die Zufallsvariable $n^{1/\alpha}(M_n - 1)$ in Verteilung gegen Ψ_α .

Beweis. Sei $t \leq 0$ beliebig. Mit Satz 1.1 erhalten wir

$$\mathbb{P}[n^{1/\alpha}(M_n - 1) \leq t] = \mathbb{P}[M_n \leq 1 + tn^{-1/\alpha}] = F^n(1 + tn^{-1/\alpha}).$$

Aus der Form der Verteilungsfunktion F folgt, dass

$$\mathbb{P}[n^{1/\alpha}(M_n - 1) \leq t] = (1 - (1 - (1 + tn^{-1/\alpha}))^\alpha)^n = \left(1 - \frac{(-t)^\alpha}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{-(-t)^\alpha}.$$

Für $t \geq 0$ gilt $\mathbb{P}[n^{1/\alpha}(M_n - 1) \leq t] = 1$. Die Behauptung folgt. \square

Definition 1.14. Die Verteilungsfunktion Ψ_α in (1.4) heißt **Weibull-Verteilung** mit Parameter $\alpha > 0$.

Bemerkung 1.15. Man kann Satz 1.13 wie folgt interpretieren: für großes n nähert sich das Maximum M_n dem Wert 1 von unten an. Die Differenz $M_n - 1$ nimmt sehr kleine negative Werte auf der Skala $n^{-1/\alpha}$ an. Reskaliert man $(M_n - 1)$ mit dem Faktor $n^{1/\alpha}$, so erhält man approximativ Weibull-verteilte Werte.

Bemerkung 1.16. Eine Fréchet-verteilte Zufallsvariable nimmt nur positive Werte an. Eine Weibull-verteilte Zufallsvariable nimmt nur negative Werte an. Eine Gumbel-verteilte Zufallsvariable kann sowohl positive als auch negative Werte annehmen.

Bemerkung 1.17. Wir haben drei verschiedene Typen von Extremwertverteilungen kennen gelernt: Gumbel, Fréchet und Weibull. Später werden wir zeigen, dass jede Extremwertverteilung sich auf eine dieser drei Verteilungen zurückführen lässt.