TECHNISCHE UNIVERSITÄT WIEN

Extremwerttheorie

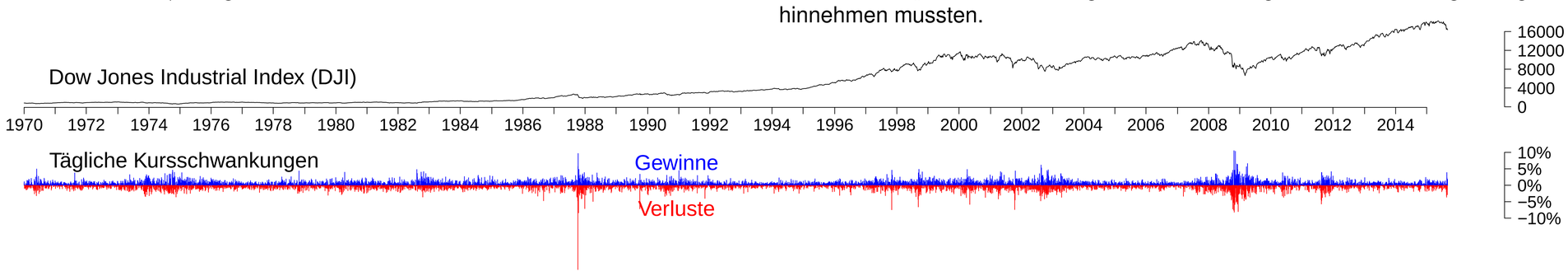
Die Mathematik der seltenen Ereignisse

Extremwerttheorie - Was ist das?

Für gewöhnlich werden aus einer größeren Stichprobe all jene Werte entfernt, die so groß oder klein sind, dass sie unplausibel erscheinen. Sie werden als Ausreißer deklariert und vom weiteren Prozess des Schätzens und Berechnens ausgeschlossen. Genau mit diesen Werten beschäftigt sich als Zweig der Statistik die Extremwerttheorie. Sie beschreibt, wie groß Ausreißer sind und mit welchen Wahrscheinlichkeiten Extrema einer vorgegebenen Höhe vorkommen. So kann man auch sehr seltene (z. B. nur alle 100 Jahre vorkommende) Ereignisse beschreiben und deren Wahrscheinlichkeiten abschätzen.

Ein Beispiel - Wann kommt der nächste Börsencrash?

Gefürchtete extreme Ereignisse sind beispielsweise Börsen-Crashs – große Kursstürze, die innerhalb von kurzer Zeit erfolgen. Die unten stehende Grafik zeigt die Entwicklung des Dow Jones Industrial Index (DJI), des wichtigsten US-amerikanischen Aktienindex, über die letzten 45 Jahre. Während der obere Teil des Bildes die Gesamtentwicklung des DJI zeigt, stellt der untere Teil die Tagesrenditen (daily returns) in Prozent dar. Aus dieser Grafik sticht ein großer Ausschlag am 19. Oktober 1987 heraus – der „Schwarze Montag“ – an dem Anleger an einem einzigen Tag Verluste von über 20% hinnehmen mussten.



Ein genauerer Blick auf die echten Daten

Um die Kursstürze mit Hilfe der Extremwerttheorie zu modellieren, müssen wir uns zunächst auf ein extremes Ereignis festlegen, welches wir betrachten wollen. In diesem Fall wählen wir als extreme Ereignisse den größten Tagesgewinn und den größten Tagesverlust, die Anleger im Zeitraum eines Monats hinnehmen müssen. Die folgenden Bilder zeigen die empirische Verteilung der Ereignisse als Histogramm basierend auf 430 Monate Beobachtung (rosarot und hellblau) und die Dichte einer angepassten Extremwertverteilung (dunkelblau).

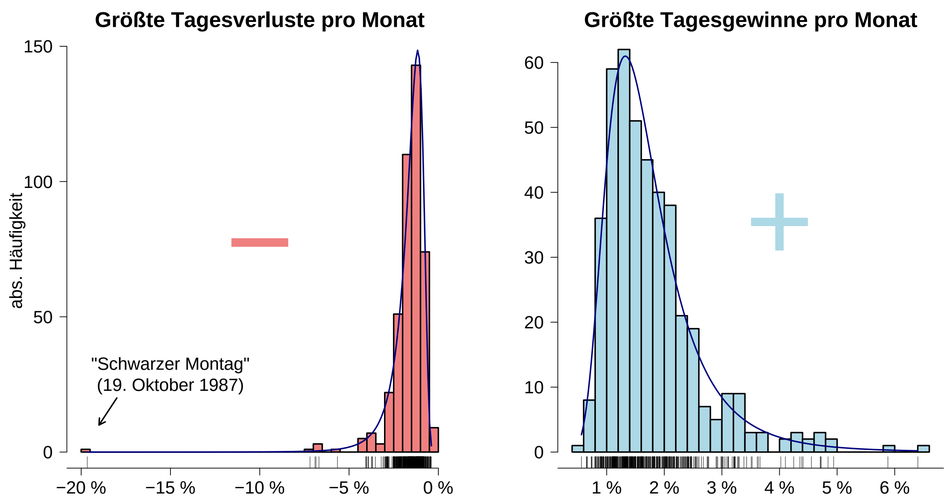


Abbildung 1: Maximaler Tagesverlust und maximaler Tagesgewinn, den Anleger innerhalb eines Monats hinnehmen mussten.

Man beachte, dass bei der linken Grafik die horizontale Skala von 0% bis -20% reicht, während sie beim rechten Teil der Abbildung lediglich bis ca. 6,5% reicht. Das zeigt, dass starke Kurssprünge nach unten häufiger Auftreten als gleichstarke Sprünge nach oben – in der Sprache der Finanzmathematik ausgedrückt ist das linke Ende der Verteilung (left tail) dicker als das rechte Ende (right tail). Das Modell der Extremwerttheorie zeigt gute Übereinstimmung mit den beobachteten Daten und erlaubt es beispielsweise Aussagen über die Wahrscheinlichkeit eines Tagesverlusts von mehr als 5% im nächsten Monat zu machen.

Die Normalverteilung - ein zentraler Fall

Wir wollen eine Münze werfen. Ergibt der Wurf „Kopf“, so sei das Ergebnis 0, bei „Zahl“ gleich 1. Wirft man nun die Münze drei Mal, so gibt es $2^3 = 8$ mögliche Ergebnisse, wir wollen aber die Summe betrachten, hier gibt es nur vier (0, 1, 2 und 3) mögliche Ausgänge mit jeweils unterschiedlichen Wahrscheinlichkeiten. Eine Binomialverteilung beschreibt diese Wahrscheinlichkeiten und damit dieses Zufallsexperiment. Eine offensichtliche Verallgemeinerung ist, nicht gerade drei, sondern viel allgemeiner, n Münzen zu werfen.
Ein sehr wichtiger Satz – der vielleicht noch aus der Schule bekannt ist – beschreibt das Verhalten der Binomialverteilung und fast aller anderen Verteilungen: Der zentrale Grenzwertsatz besagt, dass mit wachsendem n sich die (skalierte) Summe an eine Normalverteilung annähert:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \le z\right) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx = \Phi(z) \qquad \forall z \in \mathbb{R}$$

wobei hier X_i den i-ten Münzwurf bezeichnet, $\mu = \frac{1}{2}$ dessen Mittel, $\sigma = \frac{1}{2}$ die Standardabweichung, $\mathbb{P}(\mathcal{A})$ die Wahrscheinlichkeit, dass $\mathcal{A}$ eintritt, und Φ die Verteilungsfunktion der Standardnormalverteilung.

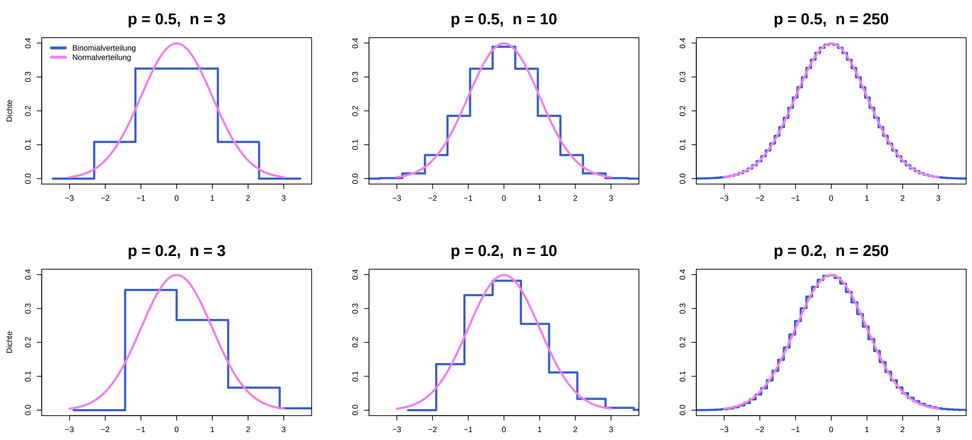


Abbildung 2: Annäherung der Binomialverteilung an eine Normalverteilung: oben für eine faire Münze: $\mathbb{P}(\text{Kopf}) = \frac{1}{2}$, und unten für eine unfaire: $\mathbb{P}(\text{Kopf}) = \frac{1}{5}$; für jeweils 3, 10 und 250 Münzwürfe.

Extremwertverteilungen - Die Mathematik dahinter

Die Statistik beschäftigt sich nicht nur mit dem klassischen Fall der Normalverteilung, sondern mit vielen anderen, für unterschiedliche Anwendungen sehr wichtigen Verteilungen. Für die Extremwerttheorie besonders bedeutend sind die Gumbel-, Fréchet- und Weibull-Verteilung.

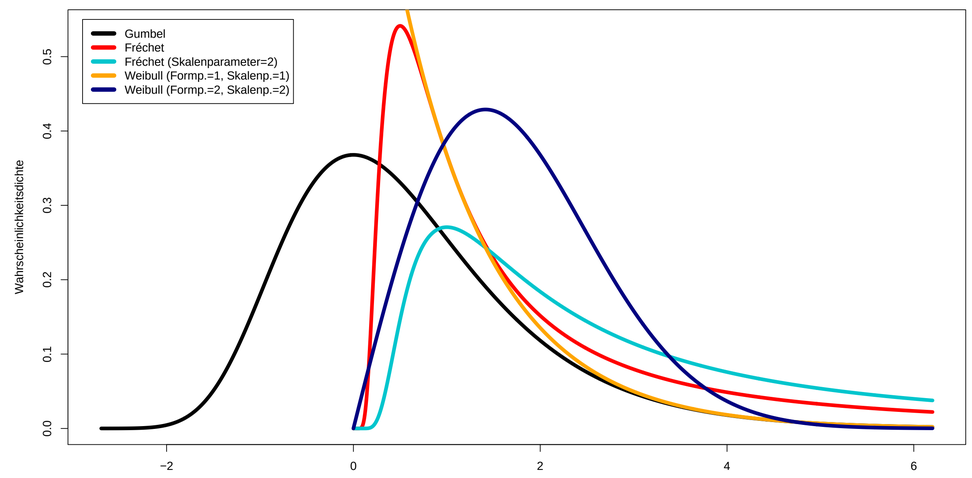


Abbildung 3: Dichten von Gumbel-, Fréchet- und Weibull-verteilten Zufallsgrößen mit teilweise unterschiedlichen Parametern.

Eine zentrale Aufgabe der Extremwerttheorie ist nun, das Maximum von n verschiedenen, unabhängigen Zufallsgrößen zu beschreiben. Das können Tagesrenditen des Aktienmarktes oder auch Wasserstände der Donau sein.
Die geplante Staumauer soll mit 95%-iger Wahrscheinlichkeit allen Fluten des nächsten Jahrhunderts stand halten. Wie hoch muss sie gebaut werden? Wie wahrscheinlich ist ein Börsencrash von mehr als 15% für das kommende Jahr? Oder, um es in die Sprache der Mathematik zu übersetzen, wie ist das Maximum aller Jahreshöchststände der Donau verteilt? Wie ist der maximale Tagesverlust des kommenden Monats (also von 21 Börsentagen) verteilt?
Es soll das Maximum oder Minimum von Zufallsgrößen untersucht und beschrieben werden. In den 20er Jahren des vergangenen Jahrhunderts haben Fisher und Tippet herausgefunden, dass für alle möglichen Verteilungen die skalierten Extrema auf nur drei verschiedene Arten verteilt sein können: Gumbel, Fréchet oder Weibull. Inzwischen ist sogar bekannt, dass diese drei Verteilungen in eine gemeinsame Klasse gehören, d. h. sich auf die selbe Art parametrisieren lassen. Der von den Nazis vertriebene deutsch-jüdische Mathematiker Emil Gumbel hat hier in den 40er Jahren für viele Anwendungen wichtige Entdeckungen gemacht.

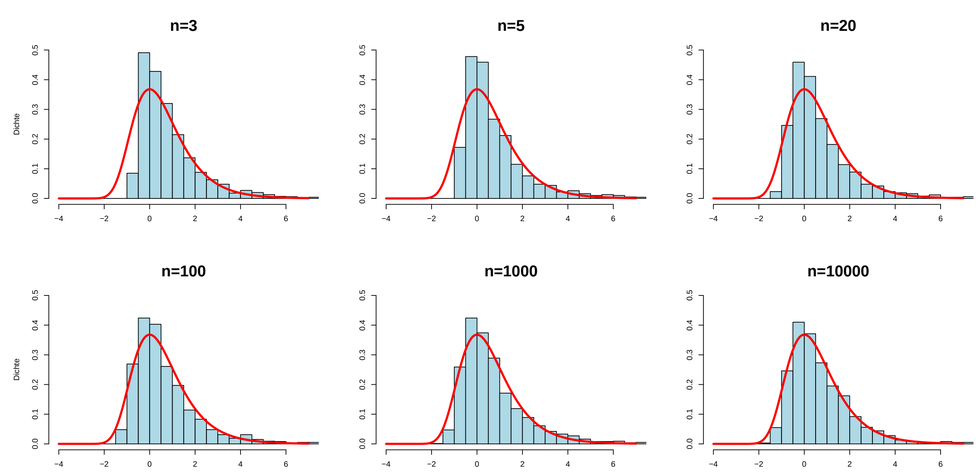


Abbildung 4: Annäherung der Verteilung des skalierten Maximums von n log-normalverteilten Zufallsgrößen an eine Gumbel-Verteilung; Stichprobenumfang: 2000.

Eine Beschreibung des Extremums von genau 12 oder 21 Zufallsgrößen ist weder praktikabel noch stets in geschlossener Form möglich. Vielmehr versucht man – genau wie beim zentralen Grenzwertsatz – jene Verteilung anzugeben, der die Verteilung der Extrema einer wachsenden Anzahl von Zufallsgrößen immer näher kommt. So findet man für z. B. unabhängige normal, log-normal, exponential oder Gamma-verteilte X_i folgende Formel:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\max\{X_1,\dots,X_n\} - b_n}{a_n} \le z\right) = \exp\{-\exp(-z)\} \qquad \forall z \in \mathbb{R}$$

wobei a_n und b_n Konstanten sind, die durch die Verteilung der X_i festgelegt sind.
In Abb. 4 ist genau diese Konvergenz zu sehen: Mit wachsender Stichprobengröße, aus der geschöpft wird, nähert sich das empirische Histogramm (hellblau) dem Grenzfall Gumbelverteilung (rot) immer besser an.

Institut für Stochastik und Wirtschaftsmathematik
Forschungsgruppe Finanz- und Versicherungsmathematik
Poster gestaltet von Martin Keller-Ressel und Thomas Steiner (2005), aktualisiert von Srećko Mihaljević (2015)