

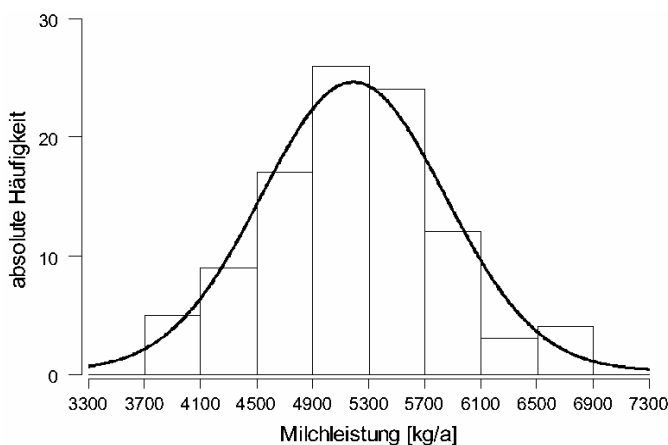
Die logarithmische Verteilung in der Natur

Die bevorzugte Verteilung in der Natur ist nicht linear sondern in erster Näherung logarithmisch.

Von Dipl.-Ing. FH André Waser

Beim Recherchieren über die statistische Verteilung verschiedener natürlicher Größen hat sich herausgestellt, daß die Natur nicht der allseits bekannten Gauß'schen Normalverteilung sondern einer logarithmischen Normalverteilung folgt. Diese Tatsache ist den Spezialisten der Statistik und Stochastik schon länger bekannt. Andererseits hören Naturwissenschaftler und Ingenieure von dieser natürlichen Verteilung an den Lehranstalten oft wenig. Die Vorstellung der logarithmischen Normalverteilung mit vielen Beispielen und einem möglichst einfachen Erklärungsversuch erscheint deshalb wertvoll.

Statistische Auswertungen von beliebigen Meßgrößen finden breite Anwendung in Naturwissenschaft und Technik. Häufig geht es dabei um die Ermittlung eines Mittelwertes aus einer genügend großen Anzahl von Meßwerten und um die Verteilung dieser gemessenen Werte um diesen Mittelwert. Wie bekannt ergibt sich in vielen Fällen eine sogenannte Gauß'sche Glockenkurve (Bild unten). Diese Kurve entspricht



Die glockenförmige Häufigkeitsverteilung der jährlichen Milchmenge pro Kuh, ermittelt über eine große Anzahl Kühen. Die häufigste Milchleistung liegt zwischen 4900 und 5300 Kilogramm. Beidseits dieses häufigsten Bereichs nimmt die Milchleistung ‚symmetrisch‘ ab.

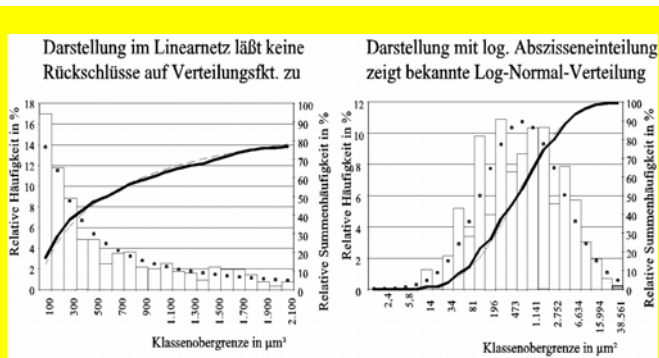


den Erwartungswerten, die sich aus der Streuung vieler Meßwerte eines bestimmten Parameters ergibt. Bei der Qualitätsüberwachung in der Produktion von Meßgeräten – zum Beispiel einer Waage, einer Benzinzapfsäule oder eines Elektrizitätszählers – werden die Produktionsdaten wie beispielsweise Kalibrierwerte laufend erfasst und statistisch ausgewertet. Entspricht die Verteilung dieser zur Eichung benötigten Kalibrierwerte der erwarteten Gauß'schen Glockenkurve, so ist alles in Ordnung. Stellt sich jedoch eine starke Unsymmetrie dieser Kurve ein, so weiß der Prozeßverantwortliche in der Produktion, daß irgendwo ein systematischer Fehler (verkoppelte Parameter) vorliegen muß, den es nun zu suchen und zu beheben gilt.

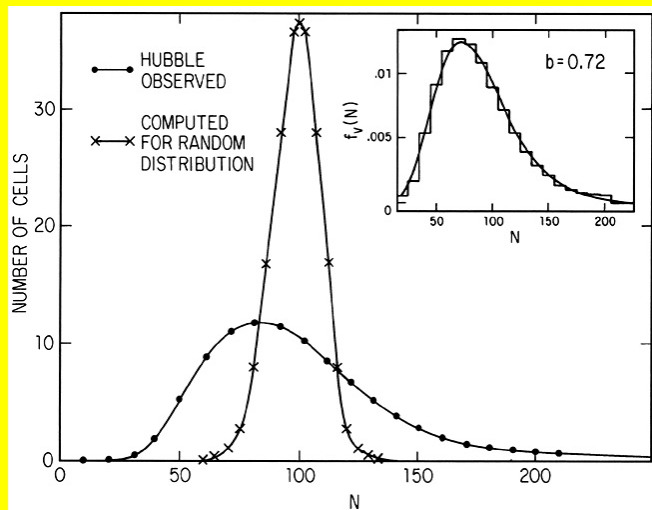
Dieses oder ähnliche Vorgehen werden auch bei der Fehlerrechnung, Optimierung, Prognose oder Qualitätsüberwachung eingesetzt und haben sich überall gut bewährt. Man neigt deshalb schnell zur Annahme, diese Gauß'sche Glockenkurve gilt für die Verteilung aller natürlichen Größen. Tatsächlich ist dies in der Natur eher selten der Fall. Hier findet man sehr oft eine unsymmetrische, nach links verschobene Verteilung um den häufigsten Wert (siehe Bilder). Statt einer linearen Gauß-Verteilung wird zur Modellierung jetzt die sogenannte Lognormalverteilung verwendet.

Nachfolgend sind Beispiele für solche logarithmische Normalverteilungen genannt, die schon mit einer Internet-Recherche gefunden werden können:

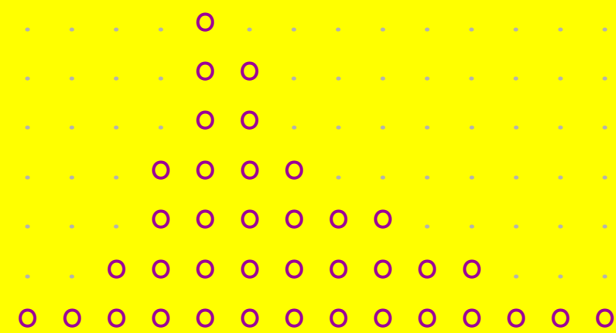
- Durchmesser von Bäumen
- Breite von Hanfbastfasern
- Durchmesser von Bakterien
- Körpergröße des Menschen
- Verteilung der Galaxien
- Dramaturgie in Filmen
- Größe von Ölfeldern
- Verteilung chemischer Elemente in Sedimenten
- Größe von Städten, Dörfern, Gemeinden usw.
- Partikelgrößen in Eis- oder Wasserwolken
- Rußpartikelgröße in Rauchgasen
- Fehlstellengrößen im Beton
- Wachstumsgröße von Kristallen
- Korngrößenverteilung nach versch. Mühlverfahren
- Täglicher Wasserdurchfluß in Flüssen
- Windgeschwindigkeiten
- Lebenszyklus einer biologischen Art
- Anzahl Angestellte in Firmen
- Bilanzgröße amerikanischer Banken
- Dichte natürlicher Fallholzlichtungen
- Für eine Heilung notwendige Penizilindosis
- Wachstum der Rißbildung bei Materialermüdung
- Korrosion von Metallen
- Häufigkeit der besuchten unterschiedlichen Seiten pro Internetsurfer.



Die Verteilung von Hanfbastfasern auf linearer (links) und logarithmischer Abszisse (rechts) nach Jörg Müssing & Hans Mehlich. Auf der linearen Abszisse vermutet man auf den ersten Blick eine hyperbolische Verteilung. Erst bei Verwendung der logarithmischen Abszisse zeigt sich die Normalverteilung (Quelle www.nova-institut.de).



Die Verteilung von Galaxien nach William C. Saslaw. Die gekreuzte Kurve zeigt die erwartete Verteilung, die gepunktete Kurve zeigt die von Erwin Hubble 1933 gemessene Verteilung (Quelle: Cambridge University Press, 1999).



Durchschnittliche Anzahl Schüsse in aufeinanderfolgenden Sekunden einer Kampfszene aus Filmen nach 1916 (Tonfilme) nach Herbert Birett. (Quelle: „Quellen zur Filmgeschichte“, www.unibw-muenchen.de/).

Diese Auflistung könnte noch stark erweitert werden. Die Lognormalverteilung wird für verschiedenste Anwendungen laufend verwendet (siehe auch Bilder links). Allerdings wird sie nur deshalb verwendet, weil es einer empirischen Erfahrung für die jeweilige Problemstellung entspricht. Was ist die dahinter liegende Ursache? Warum folgen so viele Prozesse dieser logarithmischen statt der standard Normalverteilung?

Es scheint fast, also ob die Natur nicht linear sondern logarithmisch skaliert ist. Doch vor einer vorschnellen Generalisierung sei hier schon gewarnt. Ist jetzt beispielsweise die Behauptung zulässig, daß grundsätzlich alle natürlichen Verteilungen ohne Ausnahme lognormal sind? Wohl kaum. Denn warum sollen einerseits alle oben aufgelisteten Prozesse der lognormalen Verteilung unterliegen, wenn doch das Beispiel der Milchleistung von Kühen der standard Normalverteilung folgt? Und die Milchleistung der Kuh ist trotz Mechanisierung in der Landwirtschaft (noch) ein biologischer Vorgang geblieben.

Aus obiger Liste verschiedenster Prozesse kann trotzdem eine Gemeinsamkeit zumindest erahnt werden: Es handelt sich durchwegs um Wachstums- oder Zerfallsprozesse. Diese beinhalten ein multiplikatives Gestaltungsprinzip. Das zeigt sich schon beim Wachstum eines Organismus. Alle Teile des Organismus wachsen um einen bestimmten Wachstumsfaktor, ohne die individuelle Struktur des Organismus zu verändern. Betrachtet man hingegen das Wachstum von einzelnen Organen, so findet man dort oft das additive Wachstum, also das Wachstum durch Hinzufügen neuer Zellen.

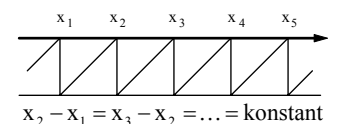
In statistischen Auswertungen werden Verschiedenheiten ausgewertet. Diese Verschiedenheiten erhält man durch Vergleich. Nun gibt es aber mindestens zwei Arten des Vergleichens.

Entweder betrachtet man die Verschiedenheiten zwischen zwei Größen als Differenz (Unterschied) oder als Quotient (Verhältnis). Beim ersten steht im Hintergrund das additive Gestaltungsprinzip, beim zweiten das multiplikative Prinzip. Beide sind in der Natur anzutreffen, oft sogar gemischt.

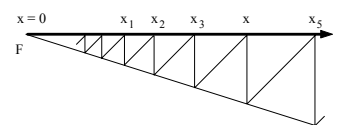
In unserem Sprachgebrauch haben wir uns längst an diese beiden Prinzipien gewöhnt. Reden wir zum Beispiel von *Verlust*, *Gewinn*, *Zuschuß*, *Abzug* werden die Unterschiede des positiven oder negativen Hinzufügens angesprochen. Reden wir von *Zins*, *Rendite*, *Schwellung*, *Schrumpfung*, dann sind die Unterschiede durch Multiplikation oder Division gemeint.

Diese zwei Gestaltungsprinzipien haben ihre Entsprechungen auch in der Mathematik. Im einen Fall spricht man von der arithmetischen Reihe, oder auch dem arithmetischen Mittel. Im anderen Fall spricht man von der geometrischen Reihe oder dem geometrischen Mittel. Tatsächlich liegen die Unterschiede zwischen der Gauß'schen Normalverteilung und der Lognormalverteilung in diesen zwei Gestaltungsprinzipien.

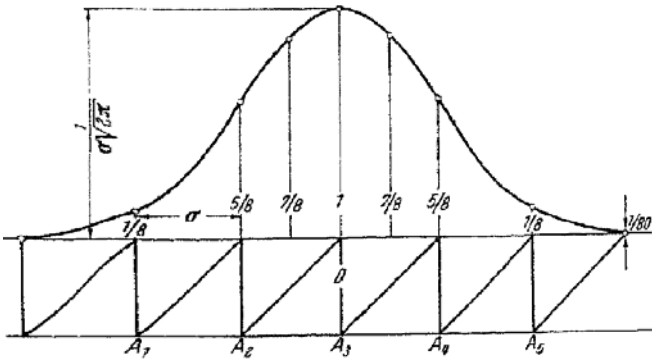
Hans Gebelein und Hans-J. Heite haben bereits 1950 in ihrer Arbeit „Über die Unsymmetrie biologischer Häufigkeitsverteilungen“ eine anschauliche Beschreibung dieser zwei Gestaltungsprinzipien aufgezeigt. Dazu bietet sich die grafische Darstellung der **arithmetischen Reihe** (konstante Abstände)



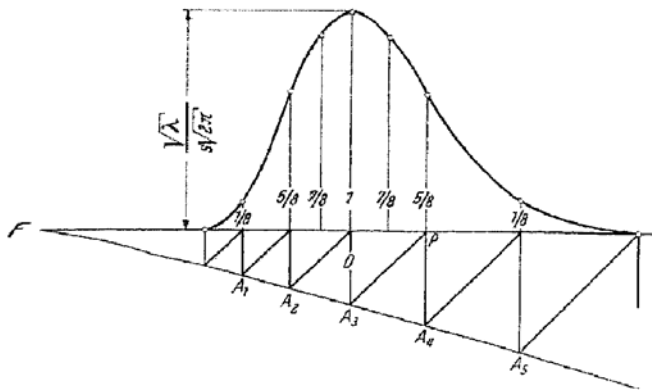
und der **geometrischen Reihe**



$\frac{x_2}{x_1} = \frac{x_3}{x_2} = \dots = \text{konstant}$
an.



Der Zusammenhang der Gauß'schen standard Normalverteilung mit den festen Abständen einer arithmetischen Reihe nach Gebelein & Heite (1950).



Die Kompression der Glockenkurve zu der logarithmischen Normalverteilung verursacht durch die geometrische Reihe. Quelle Gebelein & Heite, 1950.

Wenn der Fluchtpunkt F unendlich fern nach links verschoben wird, entsteht aus der geometrischen Reihe eine arithmetische Reihe. Diese Tatsache ist aus der projektiven Geometrie bekannt.

Durch die oben grafisch dargestellte exponentielle Komprimierung des Maßstabes der geometrischen Reihe erfolgt links eine Komprimierung und rechts eine Dekomprimierung der Gauß'schen Kurve auf der linearen Abszisse. Diese Asymmetrie wird bei der Lognormalverteilung durch Anwenden des Logarithmus (die Umkehrfunktion der in den Wachstums- und Zerfallsprozessen allgegenwärtigen Exponentialfunktion) aufgelöst und es entsteht auf der logarithmischen Skala wieder eine symmetrische Verteilung. Der Logarithmus eines Produktes zweier Zahlen ist die Summe der Logarithmen beider Zahlen. Das ist aus der Zeit der Rechen-

schieber bekannt. In Analogie dazu läßt sich verstehen, warum eine geometrische (also multiplikativ gebildete) Reihe durch Logarithmieren in eine arithmetische (also additive) Reihe transformiert wird. Und genau das geschieht bei der Verwendung der Lognormalverteilung an Stelle der Gauß'schen Normalverteilung.

Weil natürliches Wachstum zumindest teilweise dem multiplikativen Gestaltungsprinzip unterliegt, ist eine lognormale Verteilung natürlicher Werte zu erwarten.

Die logarithmische Skala erscheint unter diesem Betrachtungswinkel ganz von selbst als die natürliche Skala. So haben sich beispielsweise auch unsere Sinne (Gehör-, Sehschärfeempfinden) dieser Skala angepaßt.

Zwei weitere sehr bekannte Verteilungen sind aus der Ther-

modynamik bekannt: Die Maxwell-Boltzmann'sche Geschwindigkeitsverteilung von Gasteilchen und das Plank'sche Strahlungsgesetz. Die Maxwell-Boltzmann'sche Verteilung basiert auf vollständig elastischen Stößen zwischen den Gasteilchen in einem begrenzten (meist kugelförmigen) Raum. Dabei wird diesem Raum weder Energie zu- noch abgeführt. Dieser Prozeß hat augenscheinlich wenig mit Wachstum und Zerfall zu tun. Deshalb tritt erwartungsgemäß eine Geschwindigkeitsverteilung auf, die sich mit der standard Normalverteilung modellieren läßt. Deutlich anders sieht das bei der Plank'schen Strahlungsformel eines schwarzen Körpers aus. Besonders bekannt in diesem Zusammenhang ist die kosmische Hintergrundstrahlung mit einem Maximum bei 2.725 Kelvin. Der Intensitätsverlauf einer Strahlung kann vereinfacht als Häufigkeit der Anzahl pro Sekunde emittierten Photonen (Lichtteilchen) pro Frequenz aufgefaßt werden. Dieser Häufigkeitsverlauf ist nicht mehr symmetrisch um den Maximalwert, sondern gleicht in der Form stark einer Lognormalverteilung.

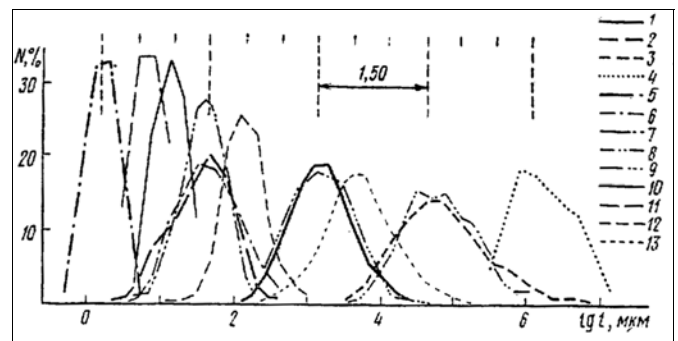
Tatsächlich läßt sich der Verlauf der kosmischen Hintergrundstrahlung näherungsweise mit einer logarithmischen Normalverteilung vergleichen. Ob dies

nur ein Zufall ist oder eine tiefergehende Bedeutung hat, entzieht sich derzeit meinen Kenntnissen.

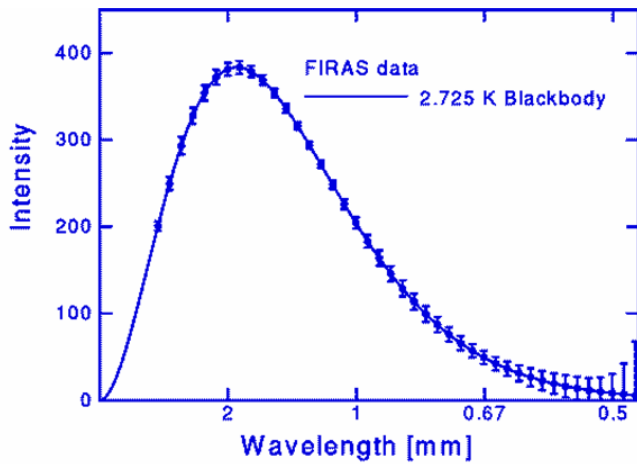
Nach dieser Spekulation ist vorerst alles Wesentliche über den Zusammenhang zwischen multiplikativen Wachstumsprozessen und deren lognormalen Verteilung gesagt. Doch ist das auch schon alles? Man könnte nun die Frage stellen, wie sich denn die Summe dieser logarithmischen Verteilungen pro Anwendung – beispielsweise die bekannten logarithmischen Verteilungen aller bekannten Fließgeschwindigkeiten von Feststoffen (zum Beispiel Glas, Eis, etc.), Flüssigkeiten, Gasen und verschiedenen Plasmen – zueinander beziehen.

Um die Prognose gleich vorweg zu nehmen: Diese einzelnen lognormalen Verteilungen werden sich in regelmäßigen Abständen auf der logarithmischen Skala anordnen. Warum? Dazu einige Beispiele:

Beispiel 1: Der Biologe Čislenko fand nach ausgiebigen Meßreihen bezüglich Körpergröße, Gewicht etc. verschiedenster biologischer Arten immer die bekannte Lognormalverteilung. Darüber hinaus fand er zusätzlich einen fixen logarithmischen Abstand von ungefähr 0.5 (bez. ein Vielfaches davon) bei logarithmischer Basis von 10 (entspricht etwa dem Abstand von 1 beim na-



Verteilung der Arten nach Čislenko aufgetragen auf eine logarithmische Abszisse. 1 – Krustentiere, 2 – Geißelträger, 3 – Fische, 4 – Wale und Delphine, 5 – Bakterien, 6. – Kieselalgen, 7 – Wimperntierchen, 8 – Kopffüssler, 9 – Geschwänzte Manteltiere, 10 – Strahlentierchen, 11 – Blaualgen, 12 – Hefepilze, 13 – Bauchfüßler (Schnecken) und Vielborstler (Seesterne)



Das Spektrum der kosmischen Hintergrundstrahlung (der Bereich der Punkte ist gemessen) zeigt große Ähnlichkeit mit einer lognormaler Verteilung.

türlichen Logarithmus zur Basis e) zwischen den einzelnen Maxima der Häufigkeitsverteilungen pro Art. Dies gilt völlig unabhängig davon, ob es sich nun um Nagetiere, Raubtiere, Wale, Insekten, Bakterien, Algen, Pilze usw. handelt. Zusätzlich zur Lognormalverteilung pro Art kommt eine diskrete Wiederholung derselben Struktur in immer gleichen Abständen auf der logarithmischen Gerade (siehe Bild).

Beispiel 2: Seit 1772 ist der exponentielle Zusammenhang der Planetenbahnen nach der Titusbode Regel bekannt. Wilhelm Bauer berechnete 1975, wie sich die mittleren Radien der Planetenbahnen in regelmäßigen logarithmischen Abständen von ungefähr 1 zur Basis des natürlichen Logarithmus e wiederholen. Daraus hat Norbert Hartun die bemerkenswerte Feststellung gemacht, daß sich die Planetenbahnen genau in Schwingungsknoten der Sonnen-Schallwelle auf der logarithmischen (nicht linearen!) Skala befinden. Wie bekannt ist, schwingt die Sonne mit einer Schwingung pro fünf Minuten und einer Amplitude von rund 2000 Kilometer. Diese Schwingung wird auch meßbar durch eine Modulation des Sonnen-

windes bis zu den Planeten übertragen und dürfte zudem auch im Gravitationsmedium unseres Sonnensystems überall vorhanden sein. Das gesamte Planetensystem (mit dem Asteroidengürtel zwischen Mars und Jupiter) verhält sich wie ein gigantisch schwingendes System, in dem sich auf der logarithmischen Abstandsskala eine stehende Welle herausbildet.

Wilhelm Bauer hat zudem gezeigt, daß die logarithmische Verteilung der Bahnradialen auch für die Monde von Jupiter, Saturn und Uranus vorhanden ist.

Beispiel 3: Nach menschlicher (und besonders Europäischer) Sicht haben sich die Schritte der Evolution im Abstand von ungefähr 1 des Zehnerlogarithmus (etwa 2 des natürlichen Logarithmus) beschleunigt (Siehe Kasten).

Problematisch wird diese Beschleunigung für die Menschheit neben der dazu ebenfalls exponentiell steigenden Bevölkerungsdichte durch die immer kürzeren Innovationszyklen. Diese sind bereits deutlich kürzer als eine Generation geworden und drohen immer unüberschaubarer für den Einzelnen zu werden. Bekannte Beispiele dazu sind die exponentielle Zunahme von Forschungsarbeiten

oder generell der gesamten Informationsmenge.

Bis zu diesem Punkt haben wir die Lognormalverteilung und die festen Abstände dieser Verteilungen auf einer logarithmischen Skala behandelt. Was jetzt aussteht, ist eine Verknüpfung dieser beiden Sachverhalte. Wie wir inzwischen wissen, sind die logarithmischen Verteilungen für alle möglichen voneinander unabhängigen physikalischen Wachstumsgrößen (Dimensionen) gültig. Ferner wiederholen sich die selben lognormalen Verteilungen in allen Größenordnungen mit stabilem Abstand auf der logarithmischen Gerade (bez. mit konstantem Multiplikationsfaktor auf der linearen Skala) wieder. Dieser Sachverhalt wird allgemein als Skalenvarianz bezeichnet.

Ergänzend zur natürlichen Lognormalverteilung gibt es ein erstaunliches mathematisches Gesetz, das eine solche Verteilung implizit beinhaltet.

Alles hat damit angefangen, als dem Mathematiker Simon Newcomb 1881 auffiel, wie die ersten paar Seiten eines Buches mit Logarithmentafeln aus der Universitätsbibliothek viel mehr Gebrauchsspuren aufzeigen als die später folgenden Seiten. Bei einem schwer zu lesenden Buch wäre das noch zu verstehen, da ein Leser nach ein paar Seiten das für ihn zu schwere Buch wieder zurückstellt. Eine Logarithmentafel ist aber nur eine sehr lange Tabelle von Zahlen mit deren Logarithmen in schön aufsteigender Folge. Ganz offensichtlich wurden von Zahlen, deren erste Ziffer eine Eins oder eine Zwei ist, viel mehr Logarithmen nachgeschlagen als für die anderen Zahlen.

Nach Newcomb's Veröffentlichung im *American Journal of Mathematics* verging fast ein halbes Jahrhundert, bis der Physiker Frank Benford – nach dem heute diese Verteilung fälschlicherweise benannt wird – die selbe Feststellung an Logarithmentafeln machte. Ben-

ford trug darauf hin an die 20'000 Zahlen aus verschiedensten Quellen für seine Publikation aus dem Jahr 1938 zusammen:

- Länge von Flüssen in USA,
- Fläche von Seen,
- Hausnummern einer Stadt,
- Alle Zahlen einer Ausgabe von Reader's Digest,
- Atomgewichte,
- Populationen,
- Preise, usw.

Im Mittelwert tritt die Eins als erste Ziffer mit einer Wahrscheinlichkeit von 30.1% auf, die Zwei mit 17.6%, die Drei mit 12.5% usw. bis zur Neun mit 4.6% (siehe Grafik unten). Benford lieferte eine Formel, mit der die Wahrscheinlichkeit P des Auftretens der ersten Ziffer d für das Dezimalsystem berechnet werden kann:

$$P(d) = \log(d+1) - \log(d)$$

Dieses Gesetz ist wirklich erstaunlich und hat das Potenzial für schlaflose Mathematiker-Nächte! Denn spontan würde jeder die Gleichverteilung annehmen (also jede erste Ziffer außer Null hat die gleiche Wahrscheinlichkeit von 11.1%).

Wichtige Ereignisse im logarithmischen Zeitraster

- Vor 5 Milliarden Jahren: Bildung des Planeten Erde.
- Vor 500 Millionen Jahre: Bildung des mehrzelligen Lebens
- Vor 50 Millionen Jahren: Vorherrschaft der Vögel und Säugetiere.
- Vor 5 Millionen Jahren: Entstehung der Anthropoiden.
- Vor 500'000 Jahren: Zeit der Homoniden.
- Vor 50'000 Jahren: Zeit der Cro Magnon – der Homo Sapiens Sapiens
- Vor 5'000 Jahren: Beginn von Zivilisierung, Siedlung etc.
- vor 500 Jahren: Bacon'sches Zeitalter: Aufkommen von Wissenschaft, Technik und Kapitalismus
- vor 50 Jahren: Vernetzung, Globalisierung, Kommunikation

Doch das trifft nicht zu. Und wer hätte das gedacht: Dieses Gesetz zur Verteilung der ersten Ziffer gilt auch für Steuererklärungen, weshalb es heute als Werkzeug für die Betrugsprävention eingesetzt wird. Ebenso kann damit in vielen Fällen eine Manipulation von Daten eindeutig festgestellt werden. Vielleicht wäre damit der Spendenskandal in Deutschland schon früher entdeckt worden. Daß sich die Kenntnis dieser Verteilung auch für Glücksspiele einsetzen läßt, zeigt der lesenswerte Aufsatz von Ian Steward im *Spektrum der Wissenschaft*.

Diese Verteilung der ersten Ziffer gilt zwar für viele Datenhaufen, aber nicht für alle. Meistens trifft diese Verteilung für Meßreihen natürlicher Werte zu. So zum Beispiel für die Halbwertszeit des radioaktiven Alpha-Zerfalls. Die Benford-Verteilung ergibt sich vorzugsweise dann, wenn die Daten einer lognormalen Verteilung unterliegen, wie Untersuchungen (zum Beispiel Scott & Fasli, 1998) gezeigt haben. Und wie wir wissen, trifft das für multiplikative Wachstumsvorgänge – und insbesondere in natürlichen Vorgängen – sehr häufig zu. Dies ist auch der Grund, warum die

Benford'sche Verteilung auch ganz exzellent bei multiplikativen Reihen wie beispielsweise der sehr bekannten Fibonacci-Reihe zutrifft.

Das Benford'sche Gesetz gilt nicht nur für dezimale Zahlensysteme sondern für alle beliebigen. Ebenso gilt sie für alle beliebigen physikalischen Größen. Mit anderen Worten: Diese Verteilung ist skaleninvariant.

Berücksichtigt man zusätzlich die Länge der Zahlen einer grossen Zahlenansammlung, so gilt für jede gleich große Zahlenlänge wieder die Benford'sche Verteilung. Im Falle des Dezimalsystems wiederholt sich dieses Benford'sche Häufigkeitsmuster im festen Abstand von Faktor Zehn, also im regelmäßigen Abstand von Eins des dekadischen Logarithmus. Bei Verwendung einer anderen logarithmischen Basis verändern sich entsprechend die Abstände, die Quantisierung aber bleibt.

In der naturnahen Technik wird oft nur die hyperbolische Substruktur betrachtet. Als wesentliches zweites Element in der naturnahen Technik muss auch die logarithmische Häufigkeitsverteilung mit einbezogen werden.

Literatur

BAUER Wilhelm, „Geo- und Astrophysik“ (1975), Selbstverlag

BENFORD Frank, „The Law of Anomalous Number“, *Proc. of the American Phil. Society* **78** (1938) 551

ČISLENKO L. L., „Die Struktur der Fauna und Flora im Zusammenhang mit den Körpergrößen der Organismen“, (1980) Verlag der Lomonosov-Universität Moskau; www.raum-energie-forschung.de/load.htm; (in Russisch)

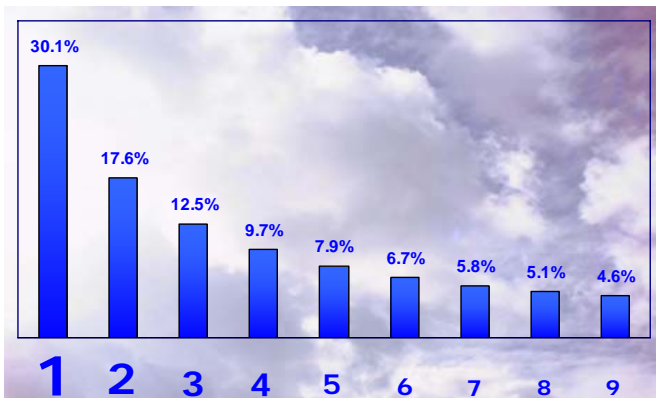
HARTHUN Norbert, „Sonne und Planeten in Resonanz – Zusammenhalt durch unbekannte Energiewelle?“ (1981), www.gruppederneuen.de

GEBELEIN Hans & Hans-J. HEITE, „Über die Unsymmetrie Biologischer Häufigkeitsverteilungen“, *Klinische Wochenschrift* **28** (15. Januar 1950) 41–45

NEWCOMB Simon, „Note on the Frequency of Use of Digits in Natural Numbers“, *American Journal of Mathematics* **4** (1881) 3940

SCOTT Paul & Maria FASLI, „Benford's Law: An Empirical Investigation and a Novel Explanation“, *University of Essex, CSM* 349 (October 2001)

STEWARD Ian, „Das Gesetz der ersten Ziffer“, *Spektrum der Wissenschaft* (April 1994) 16



Verteilung der ersten Ziffer aus Zahlen einer großen Datenmenge gemäß dem Benford'schen Gesetz.