Das „BENFORD-Gesetz“ über die Verteilung der ersten Ziffer von Zahlen

von Hans Humenberger, Wien

Zusammenfassung

Uns ist sehr wohl bewußt, daß es in der Schule nicht (oder nur kaum – evtl. in Leistungskursen mit Freiwilligen) möglich sein wird, die doch in den Anfangsgründen der Maßtheorie beheimateten Inhalte am Schluß dieses Beitrages (Betrachtungen von RAIMI) zu behandeln. Trotzdem scheinen uns auch diese einfachen maßtheoretischen Überlegungen wichtig und interessant zu sein, z.B. für Lehrerfortbildungen oder für Studierende des Mathematik-Lehramtes, selbst wenn keine wirkliche Vorbildung in dieser Richtung vorliegt. Für eine allfällige Behandlung unseres Themas in der Schule sei angemerkt, daß wir es doch für möglich und sinnvoll halten, in *Einzelfällen* (!) gewisse Ausblicke auf die Hochschulmathematik (bzw. auf wirkliche mathematische Forschungsinhalte – hier speziell der 60er-Jahre) in der Schule zu vermitteln, selbst wenn bei einem bestimmten Problem vieles nur erzählt werden kann und die dahintersteckende Mathematik eher im Hintergrund bleibt bzw. bleiben muß. Die ersten Abschnitte (Schilderung des Problems, geometrische Folgen, Beschreibung der Methode von PINKHAM und dazugehörige Kritik von RAIMI etc.) enthalten aber nur Elementarmathematik und sind insofern u.E. auch als Schulstoff in Leistungskursen denkbar. Die eigentlichen „Zielgruppen“ dieses Aufsatzes sind jedoch nicht die Schulklassen des Pflichtunterrichts, sondern die Lehrerausbildung (Mathematikstudenten), die Lehrerfortbildung (Seminare), Facharbeiten und Projekte mit interessierten Schülern etc.

1 Einleitung

Die Geschichte des BENFORD-Gesetzes begann mit Beobachtungen von Logarithmentafeln, und zwar berichtete der Physiker Frank BENFORD (1938) – nach ihm wurde das resultierende Gesetz benannt – , daß die Logarithmentafeln in den Bibliotheken auf den ersten Seiten viel dreckiger und abgegriffener wären als auf den hinteren. Dies wäre bei anderen Büchern als Logarithmentafeln in Bibliotheken durchaus erklärlich, denn viele Leute beginnen ein Buch zu lesen (Roman, Gedichte, Theaterstück, Kurzgeschichten, Sachbücher, Fachbücher etc.), hören aber vorzeitig damit wieder auf, weil sie keine Zeit mehr haben, weil es ihnen zu langweilig wird, weil es ihnen zu kompliziert wird (Fachbücher) u.ä. Wenn viele die Lektüre unfertig unterbrechen, ist es klar, daß der Anfang von Büchern abgenützter sein kann als der Schluß. Aber warum soll dies bei Logarithmentafeln der Fall

Stochastik in der Schule **16** (1996), Nr. 3, S.2-17

sein – diese werden ja nach anderen Gesichtspunkten benützt. Die einzige Erklärung, die es dafür gibt, ist, daß der Logarithmus von Zahlen mit niedrigen Anfangsziffern (1,2,...) häufiger gesucht wurde als von Zahlen mit hohen Anfangsziffern (9,8,...)! Aber warum? Kommen Zahlen mit niedrigen Anfangsziffern „in der Welt“ häufiger vor? Warum sollte die Natur eine Präferenz für die 1 als Anfangsziffer haben?

Es sind schon viele empirische Daten erhoben worden, wobei die relative Häufigkeit der einzelnen Anfangsziffern beobachtet wurde. Diese müßte bei Gleichverteilung für alle möglichen Anfangsziffern (1,2,...,8,9) bei ca. $\frac{1}{9} \approx 0,1111$ liegen – mit *Anfangsziffer* sei im folgenden stets *die erste Ziffer ungleich 0* bzw. *erste „signifikante“ Ziffer* gemeint, also z.B. 3 in 0,0367. Tatsächlich lag jedoch die relative Häufigkeit von 1 als Anfangsziffer in vielen Datensätzen (insbesondere in jenen von BENFORD) bei ca. 0,3 – abnehmend zu ca. 0,05 bei Ziffer 9.

Es sind z.B. untersucht worden:
Oberflächen vieler Seen, Hausnummern in den Adressen vieler Personen, Halbwertszeiten radioaktiver Substanzen, Energieverbrauchszahlen vieler Haushalte u.v.a.m.

Bemerkung: Es hat natürlich keinen Sinn, Daten zu betrachten, die von vornherein auf einen Bereich eingeschränkt sind, der die Möglichkeiten für die erste Ziffer ziemlich einengt – dies wären z.B. die Anzahl der Buchstaben in den Familiennamen der Bewohner einer Stadt oder eines Landes, das Alter von Studierenden an einer Universität (das Alter generell!), die Anzahl der Schulbildungsjahre, die Anzahl der Sitze in Fahrzeugen, die Wurzeln der ersten 1 000 natürlichen Zahlen usw.

Bei den meisten der Untersuchungen hat sich eine abnehmende relative Häufigkeit von 1 bis 9 als Anfangsziffer ergeben. Wenn obige Werte „tatsächlich stimmen“ (ungefähr die theoretischen Wahrscheinlichkeiten darstellen, d.h. $P(1) \approx 0,3$ und $P(9) \approx 0,05$), ist es einleuchtend, daß bei einer 9-seitigen Logarithmentafel die erste Seite abgenützter ist als die letzte (ca. sechsmal so stark!).

I. STEWART (1994, S.1) berichtet sogar von einem Jahrmarktspiel mit einem an Daten reichen Computer: Der Spieler kann selber irgendwelche Daten anklicken (z.B. Einwohnerzahlen von Verwaltungsbezirken auf den Malediven oder in Lettland, Energieverbrauch in den Ländern der Welt, und viele andere „exotische“ – jedenfalls i.a. nicht vorhersehbare Werte). Entscheidend für den Spielausgang ist die erste Ziffer der erscheinenden Zahl. Ist diese 1 oder 2, dann gehören die 10 DM Einsatz dem Spielbudenbesitzer, bei 3, 4, 5, 6, 7, 8 und 9 erhält der Spieler seine 10 DM Einsatz zurück plus 5 DM Gewinn. Auf den ersten Blick sehen die Gewinnchancen für den Spieler nicht schlecht aus, wenn man die Wahrscheinlichkeiten mit jeweils $\frac{1}{9}$ annimmt. Der Erwartungswert des Gewinnes für den Spieler beträgt dann $-10\cdot\frac{2}{9}+5\cdot\frac{7}{9}=\frac{15}{9}\approx 1,66$ DM. Wenn der Spielbudenbesitzer wirklich auf lange Sicht 1,66 DM pro Spiel bezahlen müßte, so wäre dies für ihn nicht besonders lukrativ und sein Bankrott gleichsam vorprogrammiert. Aber der Spielbudenbesitzer vertraut offenbar auf das Gesetz von BENFORD, demzufolge die Wahrscheinlichkeit für das Auftreten einer Ziffer p $(1 \leq p \leq 9)$ als erste Ziffer

$$\log_{10}(p+1) - \log_{10} p$$

beträgt. In weiterer Folge wird die Basis 10 des Logarithmus weggelassen, da nie eine andere Basis auftritt, also $\log := \log_{10}$. Danach hätten also die einzelnen Ziffern die in Tab. 1 angegebenen Wahrscheinlichkeiten (für das Auftreten als erste Ziffer).

1	2	3	4	5	6	7	8	9
0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,051	0,046

Tabelle 1: Wahrscheinlichkeiten für die einzelnen Ziffern nach BENFORD

Wenn diese Wahrscheinlichkeiten zur Berechnung des Gewinnerwartungswertes $E(G)$ obigen Spielers herangezogen werden, so ergibt sich

$$E(G) = -10\,(0,301 + 0,176) + 5\,(0,125 + 0,097 + \ldots + 0,051 + 0,046) = -2,155\ .$$

Unter dieser Voraussetzung ist das wirtschaftliche Überleben der Spielbude natürlich relativ gesichert (wenn sich genügend Kunden darauf einlassen).

2 Wahrscheinlichkeiten und Grundgesamtheit der Daten

Zunächst ist klar, daß die Wahrscheinlichkeiten für die einzelnen Ziffern, als erste Ziffer einer Zufallszahl zu stehen, von der Grundgesamtheit des „Topfes" abhängen, aus dem die Zahl zufällig gezogen wird. Wenn z.B. aus den ersten 20 natürlichen Zahlen zufällig gezogen wird, so ist offenbar die 1 als erste Ziffer ziemlich übermächtig (11 von 20 möglichen), 2 steht in zwei von 20 Fällen an erster Stelle und jede andere Ziffer $(p \neq 0)$ genau einmal!

Wir erkennen auch sofort, daß bei natürlichen Zahlen die Wahrscheinlichkeit $\frac{1}{9}$ nur dann bei jeder Ziffer $p = 1, 2, \ldots, 9$ auftritt, wenn die Grundgesamtheit aus den ersten 9, 99, 999, 9 999 usw. Zahlen besteht. Verfolgen wir z.B. einmal die Wahrscheinlichkeit von 1 als erste Ziffer, wenn der „Topf" aus den ersten n natürlichen Zahlen besteht, d.h. wir betrachten die Folge $(P_n(1))_{n\in\mathbb{N}}$:

Bei $n = 1$ ist $P(1) = 1$, bei $n = 2$ ist $P(1) = \frac{1}{2}$ usw., diese Wahrscheinlichkeit sinkt dann bis $P(1) = \frac{1}{9}$ bei $n = 9$. Dann steigt die Wahrscheinlichkeit wieder bis $n = 19$

(auf $P(1) = \frac{11}{19}$), um dann wieder bis $n = 99$ abzufallen (wieder $P(1) = \frac{1}{9}$), dann kommt natürlich wieder ein Anstieg bis $n = 199$ (auf $P(1) = \frac{111}{199}$), dem wieder ein Abfall bis $n = 999$ folgt. So geht dies natürlich „ewig" (von 10er-Potenz zu 10er-Potenz) weiter, die relative Häufigkeit wird sich nie (mit wachsendem n) bei einem stabilen Wert einpendeln, es werden nur die Phasen des Anstiegs bzw. des Abfalles klarerweise länger, die Folge $(P_n(1))_{n\in\mathbb{N}}$ ist also sicher divergent.

Einfach durch Grenzwertbildung ($\lim_{n\to\infty}$) kann man also nicht zur gesuchten Wahrscheinlichkeit $P(1)$ in ganz $\mathbb{N}$ kommen, denn $P_n(1)$ schwankt immer zwischen $\frac{1}{9}$ (Untergrenze) und einer nur fast gleichbleibenden Obergenze $(\frac{11}{19}, \frac{111}{199}, \frac{1111}{1999}, \ldots)$.

Durch fortgesetzte kumulative Mittelwertbildung – sicherlich kein Thema für den Schulunterricht – konnte z.B. B.J. FLEHINGER (1966) zeigen, daß es in einem gewissen Sinn der Wert $\log 2$ ist, um den $P_n(1)$ schwankt. Bei ihrem Modell wird der ganze Bereich aller möglichen (positiven) Konstanten der Welt (positive reelle Zahlen) durch $\mathbb{N}$ repräsentiert, da für das Problem der ersten Ziffer, die Stellung des Kommas keine Bedeutung habe und jede positive relle Zahl in Dezimalschreibweise bei Weglassen des Kommas einer natürlichen Zahl entspreche. So argumentierte FLEHINGER beim nach BENFORD benannten Gesetz des Logarithmus für alle denkbaren positiven rellen Zahlen (Konstanten).

Zur Urheberschaft bzw. Namensgebung schreibt RAIMI allerdings (1976, S.522), daß BENFORD zwar dieses Problem der Verteilung der ersten Ziffer berühmt gemacht hat, daß aber schon 57 Jahre vor ihm (1881) Simon NEWCOMB dieses „Gesetz" beschrieben hat. Es ist also in einem gewissen Sinn zu unrecht nach BENFORD benannt – ein Phänomen, das ja häufig auftritt!

3 Geometrische Folgen

Am Beispiel bestimmter geometrischer Folgen ist das BENFORD-Gesetz besonders gut und einfach zu veranschaulichen, obwohl dies natürlich noch nichts zur Klärung beiträgt, warum auch die anderen Zahlen (Konstanten) das logarithmische Gesetz befolgen sollten – vgl. RAIMI (1969b, S.110f).

Stellen wir uns eine geometrische Folge vor $(A\,q^n)_{n\in\mathbb{N}} = (A,\ A\,q,\ A\,q^2,\ A\,q^3,\ \ldots)$ mit $A, q > 0$ und eine *logarithmische* Skala, auf der wir die einzelnen Folgenglieder einzeichnen wollen. Der Abstand von 1 nach 2 ist dabei natürlich ein größerer als von 2 nach 3; der kleinste Abstand befindet sich zwischen 9 und 10 (bzw. 1) – wie z.B. bei einem Rechenstab. Auf dieser Skala denken wir uns irgendwo einen Startpunkt A, wobei wir beim Einzeichnen der weiteren Folgenglieder jeweils um $\log q$ nach rechts zu gehen haben (wenn $q > 1$, sonst nach links), denn es gilt ja $\log(A\,q^{n+1}) = \log(A\,q^n) + \log q$. Jede Multiplikation mit q „bewegt" den vorigen Punkt um ein konstantes Stück nach rechts (bzw. links), nämlich um $\log q$. Kommt man hierbei an das Ende der Skala, d.h. über 10 hinaus bzw. unter 1, so kann man – eine 10er-Potenz höher (niedriger) – die Skala rechts (links) verlassen und links (rechts) wieder „betreten". Um dieses Verlassen der Skala („Zurückspringen") zu verhindern, denken wir uns diese zu einem Kreis gebogen und die beiden Enden zusammengeschweißt („kreisförmiger Rechenstab" – siehe Abb. 1).

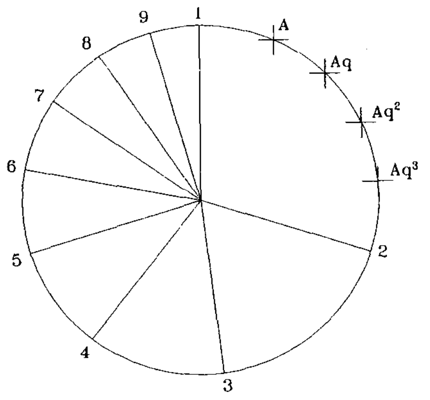

Abbildung 1: Kreisförmige logarithmische Skala

Anhand der Abb. 1 ist es nicht schwierig, sich zu überzeugen, daß viele geometrische Folgen das Gesetz von BENFORD befolgen, wenn ihre Glieder bis zu einem hinreichend großen n betrachtet werden. Der Umfang des Kreises beträgt $\log 10 = 1$ und der Abschnitt z.B. zwischen 1 und 2 (dort liegen alle Werte, die mit 1 beginnen) hat Länge $\log 2 - \log 1 = \log 2$. Damit dies wirklich die Wahrscheinlichkeit darstellt, daß ein zufällig gewähltes Folgenglied mit Ziffer 1 beginnt, muß vorausgesetzt werden, daß die Folgenglieder *gleich dicht* in den einzelnen Abschnitten zu liegen kommen, daß kein Abschnitt von den Folgengliedern „bevorzugt" bzw. „vernachlässigt" wird – eine Art „geometrische Wahrscheinlichkeit", d.h. Länge des günstigen Intervalls durch die Länge des möglichen Intervalls (des Kreises mit Umfang 1). Ist z.B. $q = 10^{\frac{1}{3}}$, so werden überhaupt nur 3 verschiedene Punkte am Kreis durch die Folge erreicht, hier kann also von einer gleichen Dichte in den Abschnitten wohl nicht die Rede sein und diese Folge wird auch das BENFORD-Gesetz nicht befolgen. Ist hingegen $q = 10^{\varepsilon}$ mit $|\varepsilon| \ll 1$, so daß bei jeder Multiplikation mit q nur um ein winziges Stück (ε) weitergerückt werden muß, so kann man sich gut vorstellen, daß (bei einer ganzen Umrundung) die Folgenglieder (Punkte) in jedem „Abschnitt" gleich dicht liegen und daher die Wahrscheinlichkeiten für die jeweilige Anfangsziffer p eines zufällig gewählten Folgengliedes der jeweiligen Bogenlänge entspricht ($\log(p+1) - \log p$, also das BENFORD-Gesetz). Wenn man nicht nur einmal und auch nicht nur einige Male den Kreis

umrundet, sondern viele Tausend oder sogar Millionen mal, so wird es auch keine Rolle spielen, wenn die Folge nicht nach einer ganzen Umrundung abgebrochen wird (bei $1\frac{1}{2}$ oder $2\frac{1}{2}$ Umrundungen wäre die Dichte in den zwei verschiedenen Kreishälften ja wohl nicht gleich!).

Genauer:

- Bei allen *rationalen* Potenzen von 10 als q (also $q = 10^{\frac{a}{b}}$ mit $a \in \mathbb{Z}$, $b \in \mathbb{N}$) wird der Ausgangspunkt A wieder erreicht (nach b Schritten bzw. a Umrundungen) und es entsteht immer wieder das gleiche „Muster". In diesem Fall nützt es daher auch nichts, nach der Rückkehr zum Ausgangspunkt, weitere Folgenglieder in den „Topf", aus dem eine Zahl zufällig gewählt wird, aufzunehmen – die Annäherung an das logarithmische Gesetz kann dadurch nicht besser werden!

- Bei *irrationalen* Potenzen von 10 muß jeweils um ein irrationales Stück weitergerückt werden, es kann daher kein Punkt zweimal „getroffen" werden, neue Punkte fallen immer zwischen zwei schon vorhandene! Weiters gilt in diesem Fall – dies ist intuitiv zwar einleuchtend, bedarf aber streng genommen eines relativ schwierigen Beweises – , daß die Dichte bei uneingeschränkter Fortführung der „Produktion" von Folgengliedern in jedem der fraglichen Abschnitte gleich groß wird, so daß die Annäherung an das logarithmische Gesetz immer besser wird.

Die unendlichen geometrischen Reihen erfüllen in den „meisten" Fällen das Gesetz von BENFORD, nämlich bei irrationalen Potenzen von 10 als Faktor – es gibt bekanntlich nur abzählbar viele rationale Zahlen im Gegensatz zu den überabzählbar vielen Irrationalzahlen (das LEBESGUE-Maß von $\mathbb{Q}$ ist Null!); bei z.B. hinreichend kleinen rationalen Werten $\left|\frac{a}{b}\right|$ wird das Gesetz ebenfalls in guter Näherung erfüllt!

Wenn also eine Zahl zufällig aus einer unendlichen geometrischen Reihe gewählt wird, ist das Zustandekommen der ungleichen Verteilung der ersten Ziffer (in den meisten Fällen) durch obige Gedanken erklärt, aber warum soll es auch dann gelten, wenn man irgendeine reelle positive Konstante wählt?

STEWART (1994, S.20) schreibt dazu: „Wir Menschen zählen in arithmetischer Folge 1, 2, 3, ... und wundern uns, ungleiche Wahrscheinlichkeiten für die Anfangsziffern zu finden. Aber das läßt sich dadurch erklären, daß die Natur mit gleichen Wahrscheinlichkeiten unter den Termen einer geometrischen Reihe wählt x, x^2, x^3, ...". Aber *warum* soll die Natur gleichsam *geometrisch* statt *arithmetisch* zählen?

Auch BENFORD selbst gab keine tiefen mathematischen Erklärungen dazu ab, sondern betrachtete es als eine Art Naturgesetz, das er und andere empirisch belegt hatten.

4 Mathematische Erklärungen

Wie schon erwähnt, hat B.J. FLEHINGER die positiven reellen Zahlen (also alle in der Natur „möglichen Konstanten“ – Zahlen) unter Vernachlässigung des Kommas durch die natürlichen Zahlen ersetzt und durch fortgesetzte kumulative Mittelwertbildung das Zustandekommen des logarithmischen Gesetzes erklärt.

4.1 Das Modell von PINKHAM

Ein anderer Mathematiker, der sich mit diesem Thema beschäftigt hat, war Roger S. PINKHAM. Er betrachtete $\mathbb{R}_0^+$ als das potentielle Universum aller möglichen physikalischen Konstanten – auf das Problem, daß einige physikalische Werte negativ sein können (z.B. Temperaturen), gehen wir hier nicht ein. Wenn nun irgendein Gesetz der Verteilung der ersten Ziffer im Reich dieser möglichen Werte existieren sollte, so kann dies laut PINKHAM nur mit einem **Invarianzprinzip** verbunden sein: Die verwendeten *Einheiten* der physikalischen Größen sind ja nicht von der Natur vorgegeben, sondern höchst willkürlich von Menschen geschaffen. Es sollte also nicht sein, daß die Verteilung der ersten Ziffer in der Welt aller möglichen physikalischen Werte davon abhängt, ob die Entfernungen in Meter oder in Zoll gemessen werden, Geschwindigkeiten in Meilen pro Stunde oder in Meter pro Sekunde usw. Umwandlungen in eine andere Einheit entsprechen i.a. Multiplikationen mit einer gewissen Zahl k. Die Werte einiger physikalischer Größen hängen aber nicht nur von der Größe der jeweiligen Einheiten ab, sondern auch von der Wahl des Nullpunktes, z.B. Temperatur. Auch die Transformation von z.B. Grad CELSIUS in FAHRENHEIT entspricht nicht einer reinen Multiplikation mit einem konstanten Faktor, sondern einer *affinen* Transformation. Beide Tatsachen machen keine großen Schwierigkeiten, sie sollen hier aber gänzlich ausgeklammert werden.

Mit der Forderung nach **Skaleninvarianz** und nach der *Existenz* einer Verteilungsfunktion $F: \mathbb{R}_0^+ \to [0, 1]$ für alle möglichen physikalischen Konstanten bewies PINKHAM, daß die einzig mögliche Verteilungsfunktion für die erste Ziffer $V(n) = \log(n+1)$ ist (Wahrscheinlichkeit, daß die erste Ziffer $\leq n$ ist, $n = 1, 2, \ldots, 9$), also BENFORDs Gesetz.

Wir möchten hier nicht auf die Einzelheiten des Beweises von PINKHAM eingehen, sondern nur auf einen Widerspruch hinweisen, den RAIMI (1969a, S.344) darin entdeckt hat. PINKHAMs Resultat wird zwar anerkannt, aber die Voraussetzung über die Existenz einer Verteilungsfunktion aller möglichen physikalischen Konstanten ist leicht zu widerlegen – dies ist ein sehr gutes Beispiel, daß ein Resultat trotz falscher Voraussetzungen richtig sein kann.

PINKHAM fordert die Existenz einer Verteilungsfunktion $F: \mathbb{R}_0^+ \to [0, 1]$ mit

1. $F(0) = 0$ und $\lim_{x\to\infty} F(x) = 1$,
2. F ist monoton wachsend und ***stetig*** (sonst hätte ein einzelner Wert eine ***positive*** Auftrittswahrscheinlichkeit $\neq 0$ bei zufälliger Wahl einer Konstanten aus $\mathbb{R}_0^+$),

wobei dann F bekanntlich ein (abzählbar additives) Wahrscheinlichkeitsmaß Φ auf $\mathbb{R}_0^+$ induziert. Sei nämlich $[a, b) \subset \mathbb{R}_0^+$ ein halboffenes Intervall, dann kann $\Phi([a, b)) := F(b) - F(a)$ als Wahrscheinlichkeit interpretiert werden, daß eine Zahl im Intervall $[a, b)$ liegt. (Üblicher in der Wahrscheinlichkeitstheorie sind eigentlich linksoffene Intervalle $(a, b]$, da aber später – bei den Mengen D_p – rechtsoffene Intervalle bzw. deren Vereinigung eine wichtige Rolle spielen werden, benützen wir diese schon hier!)

Betrachten wir zunächst einmal die Menge D_p aller positiven reellen Zahlen in Dezimalschreibweise, die eine Ziffer $\leq p$ als erste Ziffer haben. Diese Menge kann offenbar als folgende Vereinigungsmenge geschrieben werden

$$D_p = \bigcup_{n=-\infty}^{\infty} [10^n, (p+1)\,10^n) \qquad p = 1, 2, \ldots, 9 .$$

Für die Wahrscheinlichkeit $P(x \in D_p) = \Phi(D_p)$ ergibt sich daher

$$\Phi(D_p) = \sum_{n=-\infty}^{\infty} (F((p+1)\,10^n) - F(10^n)) ,$$

ein Wert, der $\log(p+1)$ ergeben soll. Nun kommt die Skaleninvarianz ins Spiel. Die wohl einfachste Art, die Skaleninvarianz zu „modellieren“, ist zu verlangen, daß

$$\Phi(D_p) = \Phi(k \cdot D_p) \qquad \forall k \in \mathbb{R}^+$$

gilt, wobei $k \cdot D_p$ die Menge aller mit k multiplizierten Elemente von D_p ist:

$$k \cdot D_p := \bigcup_{n=-\infty}^{\infty} [k \cdot 10^n, k \cdot (p+1)\,10^n) .$$

Wenn ein physikalischer Wert in einem bestimmten Einheitensystem in D_p liegt, dann wird dieser Wert, wenn eine andere Skala zugrundegelegt wird (Multiplikation mit k), in $k \cdot D_p$ liegen. Skaleninvarianz muß also bedeuten, daß die Wahrscheinlichkeit, daß eine Zahl in D_p liegt, genauso groß ist wie die Wahrscheinlichkeit, daß eine Zahl in $k \cdot D_p$ liegt; das Maß von D_p muß dem Maß von $k \cdot D_p$ gleich sein: $P(x \in D_p) = P(x \in k \cdot D_p)$ bzw. $\Phi(D_p) = \Phi(k \cdot D_p)$ für $p = 1, 2, \ldots, 9$ und $k > 0$.

Aus der geforderten Skaleninvarianz folgert aber RAIMI (als Kritik an PINKHAMs Methode) sofort, daß eine Verteilungsfunktion F aller physikalischen Konstanten, wie sie PINKHAM postuliert, nicht existieren kann. Denn wenn die geforderte Skaleninvarianz in bezug auf die erste Ziffer gelten soll, d.h. für die Mengen D_p, dann scheint es vernünftig, zu fordern, daß diese Invarianzeigenschaft auch für beliebige (meßbare) Teilmengen

$A \subset \mathbb{R}$ gelten soll, d.h. $P(x \in A) = \Phi(A) = \Phi(k \cdot A) = P(x \in k \cdot A)$ – mit welchem Recht könnte diese Eigenschaft auf die Mengen D_p eingeschränkt werden? Damit ergibt sich jedoch unmittelbar

$$F(1) = P(x \in [0,1)) \overset{\text{Skal.inv.!}}{=} P(x \in [0,k)) = F(k) \qquad \forall k > 0 \, .$$

Das hieße jedoch, daß F konstant ist, was natürlich nicht sein kann!

RAIMI führt auch noch einen philosophischen Grund gegen die Existenz einer solchen Verteilungsfunktion aller physikalischen Konstanten an, der auch dann greift, wenn man nicht bereit ist, der angesprochenen Invarianzeigenschaft über die Mengen D_p hinaus Gültigkeit zu verleihen. Wenn so eine Verteilungsfunktion F existierte, so gäbe es auch einen Wert T mit $F(T) = \frac{1}{2}$. D.h. es gäbe eine physikalische Konstante mit der Eigenschaft, daß sie genau in der Mitte des Universums aller denkbaren Konstanten aller denkbaren Welten läge. RAIMI schreibt dazu: „T takes on a mystical significance. I cannot bring myself to believe in such a number, yet this is what we must believe when we admit that scale-invariance is not to be demanded for sets like $\{x \leq T\}$." (1969a, S.344). Es wäre in der Tat auch u.E. sehr „merkwürdig", wenn ein T als *Median aller denkbaren Konstanten* existierte, so daß genau die Hälfte größer und die andere Hälfte kleiner als dieser wären.

Noch ein Argument: Wenn so eine stetige Verteilungsfunktion existierte, so gäbe es nicht nur den genannten Wert T mit $P(x \leq T) = P(0 \leq x \leq T) = \frac{1}{2}$, sondern auch ein (nicht notwendigerweise symmetrisches) endliches Intervall $I = [T - a\,,\ T + b)$ mit $a < T$, so daß $P(x \in I) = \frac{1}{2}$ ist – statt des Wertes $\frac{1}{2}$ könnte hier jeder beliebige Wert $0 < c < 1$ (z.B. 0,8) stehen. Wenn jedoch alle denkbaren Einheiten „gleichberechtigt" sein sollten, so gäbe es „viel mehr" Einheiten nahe Unendlich und nahe Null – d.h. außerhalb jedes vorgegebenen endlichen Intervalls mehr als innerhalb desselben – , so daß die numerischen Einträge in die daraus resultierenden Listen „fast alle" nahe Null bzw. Unendlich wären. Obige Aussage $P(x \in I) = \frac{1}{2}$ stünde dazu offenbar im Widerspruch.

4.2 Die Methode von RAIMI

RAIMI (1969a) selbst benützte BANACH-Maße, um die Existenz einer oben angesprochenen Verteilungsfunktion aller möglichen positiven Konstanten und eines daraus resultierenden „Mittelpunktes" all dieser zu vermeiden.

Definition: Eine auf den Teilmengen von $\mathbb{R}$ definierte reellwertige Funktion Θ heißt BANACH-Maß, wenn sie vier Bedingungen erfüllt:

(a) *Normierung*: $\Theta(\mathbb{R}) = 1$, (b) *Positivität*: $\Theta(A) \geq 0 \qquad A \subset \mathbb{R}$,

(c) *Endliche Additivität*: $A \cap B = \emptyset \;\Rightarrow\; \Theta(A \cup B) = \Theta(A) + \Theta(B)$,

(d) *Translationsinvarianz*: $\Theta(A + s) = \Theta(A) \qquad s \in \mathbb{R},\ A \subset \mathbb{R}$.

Bemerkung: (ohne Beweis, vgl. RAIMI (1969a, S.345)) Es gibt viele solcher BANACH-Maße, wobei man eine Teilmenge von $\mathbb{R}$ dann „BANACH-meßbar" nennt, wenn alle BANACH-Maße dieser Menge denselben Wert haben. Es gibt nämlich durchaus Mengen, bei denen verschiedene BANACH-Maße verschiedene Werte liefern, z.B. alle Intervalle $(-\infty, x)$ bzw. (x, ∞) – diese sind also nicht BANACH-meßbar.

Als Beispiel einer BANACH-meßbaren Menge sei folgende wichtige Menge K_a angegeben:

$$K_a := \bigcup_{n=-\infty}^{\infty} [n\,,\, n + a) \qquad \text{mit} \qquad 0 < a \leq 1 \, .$$

Diese Vereinigung halboffener Intervalle (rechtsoffen) ist in Abb. 2 dargestellt.

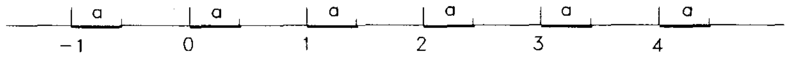

Abbildung 2: Die Menge K_a mit BANACH-Maß a

Bemerkung: Hier ist wohl auch intuitiv zu vermuten: wenn dieser Menge ein Maß zukommen soll, dann kann dies nur a sein, denn in jedem Intervall $[n, n+1)$ spiegelt sich das Verhältnis von K_a zu $\mathbb{R}$ wieder und dort ist es mit $a : 1$ abzulesen. Anders formuliert: wegen der augenscheinlichen Periodizität kann der große Bereich $\mathbb{R}$ auch *intuitiv* auf das Intervall $[0, 1)$ eingeschränkt werden. Die Wahrscheinlichkeit, daß eine „zufällig" gewählte reelle Zahl in K_a liegt, wird daher auch ohne wirkliche Exaktifizierung (z.B. durch maßtheoretische Überlegungen) mit a zu quantifizieren sein (Gleichverteilung vorausgesetzt bzw. intuitiv angenommen). Für eine allfällige Behandlung des Themas in der Schule sind also der Begriff des BANACH-Maßes und die folgenden maßtheoretischen Überlegungen (fachlich *und* methodisch) *nicht* wirklich nötig; selbst wenn bei den Schülern ein Bedürfnis nach Exaktifizierung vorhanden sein sollte, könnte dies auch anders als durch die Einführung von BANACH-Maßen geschehen (z.B. durch einen Zugang über „Packungsdichten" – vgl. ROGERS (1964), freundliche Anmerkung eines Gutachters – oder durch den am Schluß des Aufsatzes aufgezeigten Vorschlag desselben). Trotzdem seien die Überlegungen von RAIMI kurz dargestellt (insbesondere für Lehreraus- und -fortbildungen, Facharbeiten, Seminare etc.).

Es ist eine gute Übungsaufgabe für Studenten, $\Theta(K_a) = a$ ausschließlich aus den obigen vier Eigenschaften abzuleiten. Wenn wir allein aus den obigen vier Forderungen an ein BANACH-Maß zeigen können, daß das Maß von K_a den Wert a haben muß, so haben wir auch gezeigt, daß K_a überhaupt BANACH-meßbar ist. RAIMI schreibt zum Thema $\Theta(K_a) = a$ nur lapidar: „The BANACH-measure of K_a is a, as may easily be deduced from properties (a)–(d)." (1969a, S.345). Wir wollen im folgenden eine Beweisskizze geben:

Aus der endlichen Additivität (c) und aus der Translationsinvarianz (d) folgt zunächst für $a + b \leq 1$ $(a, b > 0)$

$$\Theta\left(K_{a+b}\right) = \Theta\left(K_a + (K_b + a)\right) = \Theta(K_a) + \Theta(K_b + a) = \Theta(K_a) + \Theta(K_b) \ , \qquad (1)$$

was natürlich sofort auf endlich viele Summanden verallgemeinerbar ist. Mit $K_1 = \mathbb{R}$ und (a) ergibt sich daraus unmittelbar

$$1 = \Theta(K_1) = \Theta\left(K_{\frac{1}{k}+\frac{1}{k}+\ldots+\frac{1}{k}}\right) = k \cdot \Theta\left(K_{\frac{1}{k}}\right) \qquad \forall k \in \mathbb{N} \ ,$$

woraus wir

$$\Theta\left(K_{\frac{1}{k}}\right) = \frac{1}{k} \qquad \forall k \in \mathbb{N} \qquad (2)$$

erhalten. Mit Hilfe von (1) und (2) können wir das BANACH-Maß von K_a für alle rationalen Werte von a bestimmen. Mit $a = \frac{b}{c}$ ergibt sich

$$\Theta\left(K_{\frac{b}{c}}\right) = b \cdot \Theta\left(K_{\frac{1}{c}}\right) = \frac{b}{c} \ . \qquad (3)$$

Wir haben bis jetzt also $\Theta(K_a) = a$ für rationale a gezeigt und wollen dies nun auch für irrationale Werte von a beweisen. Sei also $0 < a < 1$ eine irrationale Zahl; dann gibt es bekanntlich eine monoton nicht fallende Folge rationaler Zahlen $(a_n)_{n=1}^{\infty}$ mit $a_n \in \mathbb{Q}$ und $a_n \nearrow a$ (z.B. nach dem „Intervallschachtelungsprinzip": Folge von Intervallen mit rationalen Endpunkten $([a_n\ ,\ b_n])_{n=1}^{\infty}$, deren innerster Punkt a ist).

Mit $a_n \nearrow a$ und $b_n \searrow a$ $(a_n, b_n \in \mathbb{Q})$ gilt natürlich zunächst $a_n \leq a \leq b_n$ und somit $K_{a_n} \subset K_a \subset K_{b_n}$ bzw.

$$\Theta\left(K_{a_n}\right) \leq \Theta(K_a) \leq \Theta(K_{b_n}) \ ,$$

denn aus der Positivität (b) und aus der Additivität (c) ergibt sich allgemein:

$$A \subset B \quad \Rightarrow \quad \Theta(A) \leq \Theta(A) + \Theta(B \backslash A) = \Theta(B) \ ,$$

weil $\Theta(B \backslash A) \geq 0$ ist. Mit $\Theta(K_{a_n}) = a_n$ und $\Theta(K_{b_n}) = b_n$ ergibt sich daraus unmittelbar

$$a_n \leq \Theta(K_a) \leq b_n \qquad \forall n \in \mathbb{N} \ .$$

Da a der einzige (eindeutige) innere Punkt der Intervallschachtelung $([a_n\ ,\ b_n])_{n=1}^{\infty}$ ist, muß auch für irrationale a gelten: $\Theta(K_a) = a$.

Damit haben wir $\Theta(K_a) = a$ für alle reellen a mit $0 < a \leq 1$ bewiesen, ein Ergebnis, das später von Bedeutung werden wird.

Bemerkung: Setzt man bei der Eigenschaft (c) eines BANACH-Maßes Θ nicht nur *endliche*, sondern *abzählbare* Additivität (auch „σ-Additivität" genannt) voraus, d.h. für abzählbar viele *paarweise disjunkte* Mengen $A_n \subset \mathbb{R}$ $(n = 1, 2, \ldots, \infty)$ soll $\Theta\left(\bigcup_{n=1}^{\infty} A_n\right) =$

$\sum_{n=1}^{\infty} \Theta\left(A_n\right)$ gelten, so könnte $\Theta(K_a) = a$ für irrationale a auch folgendermaßen begründet werden:

Mit einem noch zu klärenden =-Zeichen („Stetigkeit des Maßes") erhält man:

$$\Theta(K_a) = \Theta\left(K_{\lim\limits_{n\to\infty} a_n}\right) \overset{(*)}{=} \lim_{n\to\infty} \Theta\left(K_{a_n}\right) = \lim_{n\to\infty} a_n = a \ .$$

Beweis von (*) – wir setzen formal $K_{a_0} := \emptyset$:

$$\begin{aligned}
\Theta\left(K_{\lim\limits_{n\to\infty} a_n}\right) &= \Theta\left(\bigcup_{i=1}^{\infty} K_{a_i}\right) = \Theta\left(\bigcup_{i=1}^{\infty} \left(K_{a_i} \backslash K_{a_{i-1}}\right)\right) = \sum_{i=1}^{\infty} \Theta\left(K_{a_i} \backslash K_{a_{i-1}}\right) \\
&\quad \text{(weil die Mengen } \left(K_{a_i} \backslash K_{a_{i-1}}\right) \text{ paarweise disjunkt sind!)} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \Theta\left(K_{a_i} \backslash K_{a_{i-1}}\right) = \lim_{n\to\infty} \Theta\left(\bigcup_{i=1}^{n} \left(K_{a_i} \backslash K_{a_{i-1}}\right)\right) \\
&= \lim_{n\to\infty} \Theta\left(K_{a_n}\right) \ .
\end{aligned}$$

Neben BANACH-Maßen gibt RAIMI auch den Begriff eines **skalierten Maßes** an. Die Abbildung $\log_{10} : \ \mathbb{R}^+ \to \mathbb{R}$ bildet die multiplikative Gruppe $(\mathbb{R}^+, \cdot)$ *isomorph* auf die additive Gruppe $(\mathbb{R}, +)$ ab, d.h. die Abbildung log ist bijektiv und „strukturerhaltend": $\log(a \cdot b) = \log a + \log b$. Ein **skaliertes Maß** auf $\mathbb{R}^+$ sei nun sozusagen das „Pendant" zum BANACH-Maß auf $\mathbb{R}$ – die Existenz von skalierten Maßen ist an die Existenz von BANACH-Maßen gebunden und umgekehrt. (Der Nachweis der Existenz von BANACH- und skalierten Maßen sei hier absichtlich ausgeblendet.)

Definition: Eine auf den Teilmengen von $\mathbb{R}^+$ definierte, reellwertige Funktion Ψ heißt **skaliertes Maß**, wenn sie vier Bedingungen erfüllt :

(a') *Normierung:* $\Psi(\mathbb{R}^+) = 1$, (b') *Positivität:* $\Psi(A) \geq 0 \quad A \subset \mathbb{R}^+$,

(c') *Endliche Additivität:* $A \cap B = \emptyset \quad \Rightarrow \quad \Psi(A \cup B) = \Psi(A) + \Psi(B)$,

(d') *Skaleninvarianz:* $\Psi(A\,s) = \Psi(A) \qquad s \in \mathbb{R}^+,\ A \subset \mathbb{R}^+$.

So ein skaliertes Maß Ψ kann offenbar leicht durch $\Psi(A) := \Theta(\log(A))$ gewonnen werden, wobei Θ für ein beliebiges BANACH-Maß steht. Diese skalierten Maße wirken also auf $\mathbb{R}^+$, dem potentiellen Universum aller physikalischen Konstanten – sie können auch als Wahrscheinlichkeitsmaße interpretiert werden, denn durch die ersten drei Forderungen an ein BANACH- bzw. skaliertes Maß – (a), (b), (c) bzw. (a'), (b'), (c') – sind ja die KOLMOGOROV-Axiome für eine Wahrscheinlichkeitsfunktion bzw. für ein Wahrscheinlichkeitsmaß erfüllt.

Was hat das mit dem Problem der ersten Ziffer bzw. mit dem Gesetz von BENFORD zu tun?

Diese Frage ist nun sehr leicht beantwortet! Sei nämlich analog zu oben $D_p \subset \mathbb{R}^+$ die Menge aller positiven reellen Zahlen, deren erste Ziffer in Dezimalschreibweise $\leq p$ ist. Dann läßt sich D_p – wie schon erwähnt – als folgende Vereinigungsmenge schreiben:

$$D_p = \bigcup_{n=-\infty}^{\infty} [10^n, (p+1)\,10^n) .$$

Das skalierte Maß Ψ dieser Menge D_p gibt die Wahrscheinlichkeit an, daß eine zufällig gewählte reelle Zahl eine erste Ziffer $\leq p$ hat (mit der im Maß geforderten Skalierungsinvarianz, analog zu PINKHAM). Dieses wird sich als $\log(p+1)$ herausstellen, und zwar *unabhängig* davon, *welches* BANACH-Maß zugrunde gelegt wird! Betrachten wir das schon erwähnte skalierte Maß, das über ein BANACH-Maß definiert ist: $\Psi(D_p) = \Theta(\log D_p)$. Nun gilt aber

$$\log D_p = \bigcup_{n=-\infty}^{\infty} [n, n+\log(p+1)) ,$$

d.h. die Menge $\log D_p$ ist nichts anderes als in vorheriger Diktion $K_{\log(p+1)}$. Bestimmen wir $\Psi(D_p) = \Theta(\log D_p)$, so erhalten wir – tatsächlich unabhängig von der speziellen Wahl des BANACH-Maßes Θ –

$$\Psi(D_p) = \Theta(\log D_p) = \Theta\left(K_{\log(p+1)}\right) = \log(p+1) ,$$

wie wir oben für jedes beliebige BANACH-Maß ausführlich dargestellt haben!

Bemerkungen:

- Das „dubiose" Intervall $(0,T]$, dem im Modell von PINKHAM Maß (Wahrscheinlichkeit) $\frac{1}{2}$ zukommt, ist in diesem Modell einfach *nicht meßbar*, denn $\Psi((0,T])) = \Theta((-\infty, \log T])$ und solche Mengen sind laut oben *nicht* BANACH-meßbar!
- Jedes endliche Intervall $[x,y) \subset \mathbb{R}^+$ hat 0 als skaliertes Maß, denn $\Psi([x,y)) = \Theta([\log x, \log y)) = 0$. (Jedes BANACH-Maß Θ eines endlichen Intervalles $[a,b) \subset \mathbb{R}$ muß natürlich 0 sein, weil sonst aufgrund der Translationsinvarianz ein Widerspruch zu $\Theta(\mathbb{R}) = 1$ entstünde!) Dies scheint auch philosophisch haltbar zu sein, denn wenn *alle* denkbaren physikalischen Einheiten „gleichberechtigt" sind, dann werden „fast alle" Eintragungen in den resultierenden Tabellen nahe 0 oder ∞ liegen. Salopp formuliert: Es gibt viel mehr Einheiten (und daher Zahlen) außerhalb jedes konkret vorgegebenen endlichen Intervalls als innerhalb desselben!
- Am Ergebnis und an der prinzipiellen Vorgangsweise ändert sich nichts, wenn man realistischererweise nicht $\mathbb{R}^+$, sondern $\mathbb{Q}^+$ (oder gar nur die *endlichen* Dezimalzahlen) als das mögliche Universum aller physikalischen Konstanten ansieht (bei allen in Dezimalschreibweise angegebenen Werten aller möglichen Tabellen können ja nur endlich viele Stellen berücksichtigt werden) – ohne Beweis.
- Das BENFORD-Gesetz hat mit der Darstellung im *Dezimalsystem* gar **nichts** zu tun. Auch bei Darstellungen mit jeder anderen natürlichen Zahl $a > 2$ als Basis ergäbe sich ein analoges Gesetz für die Wahrscheinlichkeit, daß eine Zahl mit einer Ziffer $\leq p$ beginnt: $P(x \in D_p) = \log_a(p+1) \quad p = 1,2,3,\ldots,a-1$.

Von einem mir unbekannten Gutachter, der freundlicherweise eine sehr ausführliche Stellungnahme zum vorliegenden Aufsatz schrieb, stammt folgende Form eines Beweises, daß das *logarithmische Gesetz* äquivalent mit der *Skaleninvarianz* ist.

Symbole:

1. Der algebraische Hintergrund der Einschränkung von $\mathbb{R}$ auf das Intervall $[0,1)$ – vgl. Abb. 2 und die darauf folgende Bemerkung – ist gegeben durch: $\widehat{[0,1)} := \mathbb{R}/\mathbb{Z}$ wird mit dem zu einem „Kreis zusammengebogenen" Intervall $[0,1)$ identifiziert. Man kann sich dazu auch vorstellen, daß die reelle Zahlengerade auf einem Kreis mit Umfang 1 „aufgewickelt" wird.
2. Die Betrachtung der *Anfangs*ziffer einer Zahl führt algebraisch auf den Übergang von $(\mathbb{R}^+,\cdot)$ zu $(\mathbb{R}^+/<10>,\cdot)$, wobei $<10>$ die (multiplikative) Untergruppe der Zehnerpotenzen ist. Die Menge $\widehat{[1,10)} := (\mathbb{R}^+/<10>)$ wird dabei mit dem zu einem „Kreis zusammengebogenen" Intervall $[1,10)$ identifiziert – vgl. Abb. 1.

Nun werden folgende **Definitionen** (etwas abgeändert) gegeben:

- Eine Zufallsgröße $X : \mathbb{R} \to \widehat{[1,10)}$ heiße BENFORD-*verteilt*, wenn $P(1 \leq X \leq a) = \log a$ für alle $a \in [1,10)$ gilt.

  Mit $Y := \log X$ hat die Zufallsgröße Y den Wertebereich $\widehat{[0,1)}$, also $Y : \mathbb{R} \to \widehat{[0,1)}$.
- Ein zur Zufallsgröße X gehöriges Wahrscheinlichkeitsmaß $\Psi : \widehat{[1,10)} \to [0,1]$, definiert durch $\Psi(A) := P(X \in A)$, heiße *skaleninvariant*, wenn $\Psi(A) = \Psi(s\,A)$ für alle Intervalle $A \subset \widehat{[1,10)}$ und alle $s \in \mathbb{R}^+$ gilt.
- Ein zur Zufallsgröße Y gehöriges Wahrscheinlichkeitsmaß $\Theta : \widehat{[0,1)} \to [0,1]$, definiert durch $\Theta(B) := P(Y \in B)$, heiße *translationsinvariant*, wenn $\Theta(B) = \Theta(B+t)$ für alle Intervalle $B \subset \widehat{[0,1)}$ und alle $t \in \mathbb{R}^+$ gilt.

Satz: Sei $X : \mathbb{R} \to \widehat{[1,10)}$ eine stetige Zufallsgröße. Dann sind die beiden folgenden Aussagen äquivalent:

1. X ist BENFORD-verteilt.

2. Ψ ist skaleninvariant.

Beweis: Die Beweisidee kann in drei Schritten dargelegt werden:

1. Durch Logarithmieren wird die Äquivalenz der Skaleninvarianz von Ψ und der Translationsinvarianz von Θ gezeigt.

2. Die Translationsinvarianz von Θ bedingt die Gleichverteilung von Y (und umgekehrt).

3. Durch Exponieren ergibt sich die Äquivalenz der Gleichverteilung von Y und der BENFORD-Verteilung von X.

Im einzelnen: Sei A ein Intervall in $\widehat{[1,10)}$. Mit $Y := \log X$, $B := \log A$ und $t := \log s$ erhalten wir:

1. Ψ ist skaleninvariant $\iff$ $\Psi(A) = \Psi(s\,A)$ $\iff$ $P(X \in A) = P(X \in s\,A)$ $\iff$ $P(\log X \in \log A) = P(\log X \in \log A + \log s)$ $\iff$ $P(Y \in B) = P(Y \in B+t)$ $\iff$ $\Theta(B) = \Theta(B+t)$ $\iff$ Θ ist translationsinvariant.

2. Θ ist translationsinvariant $\iff$ Y ist gleichverteilt: Aus der Gleichverteilung folgt klarerweise die Translationsinvarianz. Ist umgekehrt Θ translationsinvariant, folgt sofort durch Unterteilung des Einheitsintervalls in n gleichlange Teilintervalle $P(Y \leq \frac{1}{n}) = \frac{1}{n}$ für alle $n \in \mathbb{N}$. Die Additivität der Wahrscheinlichkeit bedingt dann $P(Y \leq \frac{k}{n}) = \frac{k}{n}$ für alle $k, n \in \mathbb{N}$, woraus die Gleichverteilung von Y auf allen rationalen Punkten des Einheitsintervalls folgt. Aus der Stetigkeit (oder auch aus der Monotonie der Verteilungsfunktion) folgt schließlich die Gleichverteilung auf ganz $\widehat{[0,1)}$ (auch auf den irrationalen Punkten!) – vgl. auch die Berechnung des BANACH-Maßes *Theta* der oben erwähnten Mengen K_a. Es wird hier also Ähnliches vollzogen, ohne die Begriffe BANACH-Maß bzw. *skaliertes Maß* zu strapazieren.

3. Y ist gleichverteilt auf $\widehat{[0,1)}$ $\iff$ $P(0 \leq Y \leq y) = y$ für alle $y \in \widehat{[0,1)}$ $\iff$ $P(10^0 \leq 10^Y \leq 10^y) = y$ für alle $y \in \widehat{[0,1)}$ $\iff$ $P(1 \leq X \leq x) = \log x$ für alle $x \in \widehat{[1,10)}$ $\iff$ X ist BENFORD-verteilt.

Zusammenfassung: *Wenn* es überhaupt eine Wahrscheinlichkeit für das Auftreten von Ziffern als erste Ziffer gibt (wenn also D_p ein Wahrscheinlichkeitsmaß $\Phi(D_p)$ besitzt) – und die empirischen Beobachtungen unterstützen diese These – , und wenn man vernünftigerweise zusätzlich die Unabhängigkeit von zugrunde gelegten Skalen fordert, dann muß $\Phi(D_p) = \log(p+1)$ sein! Dies erklärt also, warum die Daten von Energieverbrauchszahlen, die Oberflächen von Seen in Quadratmeilen oder in Quadratmillimeter und alle anderen physikalischen Größen, bei denen es Einheiten gibt, das BENFORD-Gesetz „befolgen". Da aber der Begriff der Skaleninvarianz bei *Hausnummern* sicher nichts zu bedeuten hat, so bleibt die dadurch unbeantwortete Frage, warum sich auch die Hausnummern in den Privatadressen z.B. aller deutschsprachigen Universitätsangestellten an dieses Gesetz halten sollten? Hier gibt wohl das Modell von FLEHINGER (natürliche Zahlen, kumulative Mittelwertbildung) eher eine akzeptable Erklärung. Ein möglicher philosophischer (*a priorischer*) Hintergrund des logarithmischen Gesetzes nach BENFORD bleibt jedoch ohnehin im Verborgenen!

Literatur:

1. BENFORD, F. (1938): The law of anomalous numbers. In: Proceedings of the American Philosophical Society **78**, 551–572.

2. BUCK, B.; A.C. MERCHANT; S.M. PEREZ (1993): An Illustration of Benford's first digit law using alpha decay half lives. In: European Journal of Physics **14**, 59-63.

3. FLEHINGER, B.J. (1966): On the probability that a random integer has initial digit a. In: American Mathematical Monthly **73**, 1056–1061.

4. KRÄMER, W. (1990): Das Gesetz der abnormalen Zahl. In: Stochastik in der Schule **10**, 3, 48–52.

5. PINKHAM, R.S. (1961): On the distribution of first significant digits. In: Annals of Mathematical Statistics **32**, 1223–1230.

6. RAIMI, R.A. (1969a): On the distribution of first significant figures. In: American Mathematical Monthly **76**, 342–348.

7. RAIMI, R.A. (1969b): The peculiar distribution of first digits. In: Scientific American **221**, Dezember 1969, 109–120.

8. RAIMI, R.A. (1976): The first digit problem. In: American Mathematical Monthly **83**, 521–538.

9. ROGERS, C.A. (1964): Packing and Covering. Cambridge.

10. STEWART, I. (1994): Mathematische Unterhaltungen. In: Spektrum der Wissenschaft (April 1994), 16–20.

Anschrift des Verfassers:

Hans HUMENBERGER, Institut für Mathematik und Angewandte Statistik, Universität für Bodenkultur, Gregor Mendel-Straße 33, A – 1180 Wien.