



FAKULTÄT FÜR INFORMATIK UND MATHEMATIK
LEHRSTUHL FÜR ANGEWANDTE MATHEMATIK
PROF. DR. BRIGITTE FORSTER-HEINLEIN

Bachelorarbeit im Fach Mathematik

Das Benfordsche Gesetz

Von abgenutzten Seiten in Logarithmenbüchern bis hin zu
einer umfassenden maßtheoretischen Analyse –
verschiedene Beweisansätze im historischen Überblick
und ein Exponat für das Mathemuseum

Verfasserin: Susanne Lack

04. Juli 2016

Zusammenfassung

Im *theoretischen Teil* dieser Arbeit wird ein Überblick über die Entwicklung der Erkenntnisse über das Benfordsche Gesetz gegeben. Im ersten Kapitel wird zunächst das Gesetz formuliert und anschließend kurz der geschichtliche Hintergrund der Entdeckung dieser Gesetzmäßigkeit erklärt. Dann wird etwas detaillierter auf die frühen Ansätze von S. Newcomb und F. Benford eingegangen, die allerdings beide noch keinen mathematischen Beweis für ihre Beobachtungen lieferten. In Kapitel 2 betrachten wir Zahlenfolgen und untersuchen, wann diese Benford-Folgen sind, d.h. wann diese dem Benfordschen Gesetz genügen. In den folgenden drei Kapiteln bewegen wir uns in \mathbb{R}^+ und wollen zeigen, dass das Gesetz von Benford gilt. Wir betrachten drei Beweise, die chronologisch, aber auch zum Teil inhaltlich aufeinander aufbauen. Nach R. Pinkham (Kapitel 3) reicht die Annahme der Skaleninvarianz und die Annahme der Existenz einer Verteilungsfunktion aller positiven Konstanten, also \mathbb{R}^+ , aus, um Benfords Gesetz zu beweisen. R. Raimi (Kapitel 4) argumentiert allerdings, dass es eine derartige Verteilungsfunktion nicht geben kann. Er selbst arbeitet mit nicht σ -additiven Maßen, dem Banach-Maß und dem skalierten Maß. Bei T. Hill (Kapitel 5) findet sich der aktuellste und umfassendste Ansatz. Nach Hill sollte ein vernünftiges Gesetz über die Verteilung von führenden Ziffern, wie das Benfordsche Gesetz, baseninvariant sein. Er führt zunächst einen geeigneten Wahrscheinlichkeitsraum ein und zeigt, dass das einzige stetige Wahrscheinlichkeitsmaß darauf das Benfordsche-Wahrscheinlichkeitsmaß ist. In Kapitel 6 wird schließlich noch ein Ausblick auf Anwendungen des Benfordschen-Gesetzes gegeben.

Im Teil *Für das Mathemuseum* wird das Exponat vorgestellt, das für das Mathemuseum der Universität Passau angefertigt wurde.

Inhaltsverzeichnis

Zusammenfassung	1
I Theoretischer Teil	4
1 Einführung	5
1.1 Das Benfordsche Gesetz	5
1.2 Historischer Hintergrund	10
1.3 Newcombs Mantissengesetz	12
1.4 Benfords Erkenntnisse	15
1.4.1 Empirische Befunde	15
1.4.2 Ordnung der natürlichen Zahlen	18
2 Benford Folgen	27
2.1 Gleichverteilung modulo 1	27
2.2 Die Verteilung von führenden Ziffern	33
3 Die Methode von Pinkham	40
4 Die Methode von Raimi	46
4.1 Kritik an Pinkham	46

4.2	Banach-Maß und skaliertes Maß	48
4.2.1	Grundlagen	48
4.2.2	Das Banach-Maß und das skalierte Maß	50
5	Die Methode von Hill	57
5.1	Grundlagen	58
5.2	Die Mantissen- σ -Algebra	62
5.3	Wahrscheinlichkeitsmaß und Baseninvarianz	71
6	Anwendungen des Benfordschen Gesetzes	81
6.1	Aufdeckung von Betrug	81
6.2	Anwendung in der Informatik	82
6.3	Überprüfung von Prognosemethoden	83
II	Für das Mathemuseum	84
	Exponat für das Mathemuseum	85
	Informationsblatt für das Mathemuseum	87
	Beispieldatensatz	89
	Literaturverzeichnis	90

Teil I

Theoretischer Teil

KAPITEL 1

Einführung

1.1 Das Benfordsche Gesetz

Das *Benfordsche Gesetz*, auch bekannt als das *Gesetz der führenden Ziffern*, besagt, dass führende Ziffern *nicht* gleichverteilt vorkommen, wie man vielleicht vermuten könnte, sondern einer speziellen logarithmischen Verteilung unterliegen (vgl. [BH15, S. 1]).

Mit *führender Ziffer* bzw. *erster signifikanter Ziffer* sei im Folgenden stets die erste Ziffer ungleich 0 gemeint, also zum Beispiel die 3 in π oder die 4 in 0,04536. (Für eine formale Definition vgl. Kapitel 5.)

Nach Benfords Gesetz tritt die 1 als führende Ziffer mit einer Wahrscheinlichkeit von mehr als 30% auf, während die 9 mit weniger als 5% Wahrscheinlichkeit vorkommt (vgl. [BH15, S. 1]).

Benfords *Gesetz für die erste signifikante Ziffer* lautet:

Theorem 1.1.1 (Gesetz für die erste signifikante Ziffer (vgl. [BH15, S. 1]))

Die Wahrscheinlichkeit P für das Auftreten der ersten signifikanten Ziffer $D_1 = d$

(im Dezimalsystem) ist gegeben durch

$$P(D_1 = d) = \log_{10} \left(1 + \frac{1}{d} \right) \quad \text{für alle } d = 1, 2, \dots, 9. \quad (1.1)$$

Das *verallgemeinerte Benfordsche Gesetz* spezifiziert nicht nur die Wahrscheinlichkeiten für das Vorkommen der ersten signifikanten Ziffer, sondern auch der zweiten, dritten und höheren signifikanten Ziffern und noch allgemeiner auch die gemeinsame Verteilung *aller* führenden Ziffern. So lässt sich damit beispielsweise die Wahrscheinlichkeit berechnen, dass eine zufällige Zahl mit der Ziffernfolge 123 beginnt. Das Benfordsche Gesetz gilt nicht nur für Zahlen im Dezimalsystem, sondern auch für jede beliebige Basis b mit $b > 1$ (dazu vergleiche Kapitel 5).

Theorem 1.1.2 (Verallg. Gesetz der führenden Ziffern (vgl. [BH15, S. 9]))

Die Wahrscheinlichkeit für die *gemeinsame Verteilung der ersten n signifikanten Ziffern* ($n \in \mathbb{N}$) ist gegeben durch

$$P(D_1^{(b)} = d_1^{(b)}, \dots, D_n^{(b)} = d_n^{(b)}) = \log_b \left(1 + \frac{1}{\sum_{i=1}^n b^{n-i} d_i^{(b)}} \right) \quad (1.2)$$

mit $d_1 \in \{1, 2, \dots, b-1\}$ und $d_j \in \{0, 1, \dots, b-1\}$ für $j = 2, \dots, n$.

$D_1^{(b)}, D_2^{(b)}, \dots, D_n^{(b)}$ bezeichnen die ersten, zweiten bzw. n -ten signifikanten Ziffern zur Basis b .

Beispiel 1.1.3

Beispielsweise ist $D_1^{(10)}(0, 345) = 3$, $D_2^{(10)}(0, 345) = 4$ und $D_3^{(10)}(0, 345) = 5$.

Beispiel 1.1.4

Die Wahrscheinlichkeit, dass eine zufällige Zahl (zur Basis $b = 10$) mit der Ziffernfolge 123 beginnt, ist gegeben durch

$$P(D_1^{(10)} = 1, D_2^{(10)} = 2, D_3^{(10)} = 3) = \log_{10} \left(1 + \frac{1}{123} \right) = \log_{10} \left(\frac{124}{123} \right) \approx 0,003517.$$

Bemerkung 1.1.5

Für $D_i^{(b)}$ wird bei $b = 10$ häufig auch kurz D_i geschrieben.

Aus der Gleichung (1.2) lässt sich unter anderem auch eine Formel für die Auftretenswahrscheinlichkeit der *zweiten signifikanten Ziffer* bestimmen.

Korollar 1.1.6 (Gesetz für die zweite signifikante Ziffer (vgl. [BH15, S. 2]))

Die Wahrscheinlichkeit P für das Auftreten der zweiten signifikanten Ziffer $D_2 = d$ (im Dezimalsystem) ist gegeben durch

$$P(D_2 = d) = \sum_{j=1}^9 \log_{10} \left(1 + \frac{1}{10j + d} \right), \quad (1.3)$$

für alle $d = 0, 1, \dots, 9$.

Diese Wahrscheinlichkeit ist ebenfalls *nicht* gleichverteilt, sondern strikt abnehmend für die Ziffern $0, 1, \dots, 9$, obwohl sie dennoch näher an der Gleichverteilung liegt als die Wahrscheinlichkeiten der ersten signifikanten Ziffern (vgl. Tabelle 1.1).

Beispiel 1.1.7

Die Wahrscheinlichkeit, dass die zweite signifikante Ziffer einer zufälligen Zahl (im Dezimalsystem) gleich 2 ist, beträgt

$$\begin{aligned} P(D_2 = 2) &= \sum_{j=1}^9 \log_{10} \left(1 + \frac{1}{10j + 2} \right) \\ &= \log_{10} \left(1 + \frac{1}{12} \right) + \log_{10} \left(1 + \frac{1}{22} \right) + \dots + \log_{10} \left(1 + \frac{1}{92} \right) \approx 0,108821. \end{aligned}$$

Tabelle 1.1 zeigt die Auftretenswahrscheinlichkeiten der ersten vier signifikanten dezimalen Ziffern in Prozent nach dem Benfordschen Gesetz, wobei die erste Zeile dem Gesetz für die erste führende Ziffer genügt.

d	0	1	2	3	4	5	6	7	8	9
$P(D_1 = d)$	0	30,10	17,60	12,49	9,69	7,91	6,69	5,79	5,11	4,57
$P(D_2 = d)$	11,96	11,38	10,88	10,43	10,03	9,66	9,33	9,03	8,75	8,49
$P(D_3 = d)$	10,17	10,13	10,09	10,05	10,01	9,97	9,94	9,90	9,86	9,82
$P(D_4 = d)$	10,01	10,01	10,00	10,00	10,00	9,99	9,99	9,99	9,98	9,98

Tabelle 1.1: Wahrscheinlichkeiten (in Prozent; abgeschnitten nach zwei Dezimalen) der ersten vier signifikanten dezimalen Ziffern nach (1.2).¹

Eine vielleicht etwas verwunderliche Folgerung aus dem allgemeinen Benfordschen Gesetz 1.1.2 ist die Tatsache, dass die signifikanten Ziffern *abhängig* voneinander sind und nicht unabhängig, wie man vielleicht annehmen könnte.

Satz 1.1.8 (vgl. [BH15, S. 2 f])

Die signifikanten Ziffern $D_1^{(b)}, D_2^{(b)}, \dots, D_n^{(b)}$, mit $n \in \mathbb{N}$, $n > 1$, sind von einander *abhängig*.

Beweis:

Wären die signifikanten Ziffern $D_1^{(b)}, \dots, D_n^{(b)}$, mit $n \in \mathbb{N}$, $n > 1$ voneinander unabhängig, so müsste für alle n und alle b gelten:

$$P(D_1^{(b)} = d_1^{(b)}, \dots, D_n^{(b)} = d_n^{(b)}) = P(D_1^{(b)} = d_1^{(b)}) \cdot \dots \cdot P(D_n^{(b)} = d_n^{(b)}).$$

Allerdings gilt beispielsweise (im Dezimalsystem)

$$P(D_1 = 1, D_2 = 2) = \log_{10} \left(1 + \frac{1}{12} \right) \approx 0,034762.$$

Aber

$$P(D_1 = 1) \cdot P(D_2 = 2) \approx 0,032759.$$

Da $P(D_1 = 1, D_2 = 2) \neq P(D_1 = 1) \cdot P(D_2 = 2)$ folgt die Behauptung. □

¹[BH15, S. 2]

Ebenfalls lassen sich mit Hilfe von Theorem 1.1.2 *bedingte Wahrscheinlichkeiten* berechnen. So beträgt zum Beispiel die Wahrscheinlichkeit, dass die zweite Ziffer eine 1 ist, unter der Bedingung, dass die erste Ziffer ebenfalls eine 1 ist

$$P(D_2 = 1 \mid D_1 = 1) = \frac{P(D_1 = 1, D_2 = 1)}{P(D_1 = 1)} = \frac{\log_{10}\left(1 + \frac{1}{11}\right)}{\log_{10} 2} = 0,125531.$$

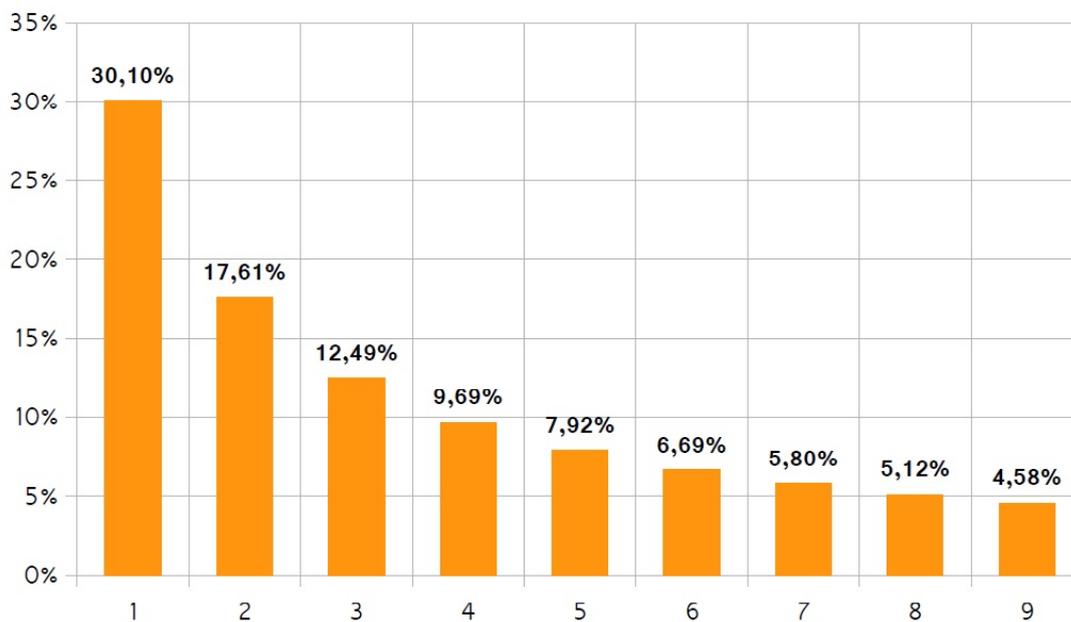


Abbildung 1.1: Benford-Wahrscheinlichkeiten für die 1. Ziffer (in Prozent, gerundet auf 2 Nachkommastellen)

Eine genauere Formulierung des Benford-Wahrscheinlichkeitsmaßes und eine Einbettung in einen geeigneten Wahrscheinlichkeitsraum erfolgt in Kapitel 5.

1.2 Historischer Hintergrund

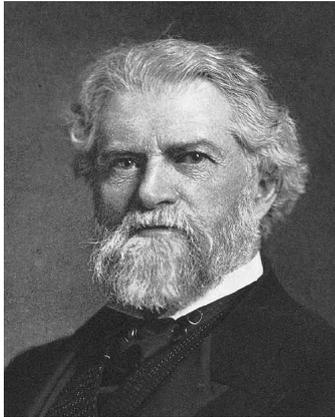
Der nachfolgende Abschnitt orientiert sich an [Hum08, S. 24 f] und [Hü07, S. 1 ff].

Die erste bekannte Abhandlung über die Verteilung führender Ziffern, geht auf den Astronomen und Mathematiker Simon Newcomb im Jahr 1881 zurück. Newcomb stellte bei der Arbeit mit Logarithmenbüchern fest, dass deren Anfangsseiten viel abgegriffener waren als die hinteren Seiten. Bei Romanen beispielsweise wäre dies durchaus verständlich, da viele Leute ein Buch zu lesen beginnen, aber oft wieder vorzeitig damit aufhören. Dadurch könnte erklärt werden, warum die vorderen Seiten abgegriffener sind. Aber warum ist dies auch bei Logarithmentafeln der Fall, denn diese werden ja bekanntlich nicht von vorne nach hinten gelesen? Nach Newcomb ist hierfür die einzige mögliche Erklärung, dass Logarithmen von Zahlen mit niedrigen Anfangsziffern, die sich in Logarithmenbüchern weiter vorne befinden, häufiger gesucht werden als die von Zahlen mit hohen Anfangsziffern. Zahlen mit niedrigen Anfangsziffern haben also eine höhere Auftretenswahrscheinlichkeit als Zahlen mit hohen Anfangsziffern. Einen mathematischen Beweis, um seine Beobachtungen zu stützen, lieferte Newcomb nicht.

57 Jahre später machte der Physiker Frank Benford im Jahr 1938 dieselbe Entdeckung mit Logarithmenbüchern. Fasziniert von diesem Phänomen sammelte er eine große Menge von Daten aus den unterschiedlichsten Bereichen und bestätigte damit die Existenz einer logarithmischen Verteilungsfunktion für die führenden Ziffern. Einen allgemeinen mathematischen Beweis lieferte allerdings auch er nicht.

Die erste mathematische Erklärung stammte von Roger S. Pinkham aus dem Jahr 1961 (vgl. Kapitel 3). Weitere Beweise und Erklärungen folgten, die zum Teil von Pinkhams Ansatz weitgehend unabhängig waren (vgl. z.B. [Fle66]), während andere den Gedanken von Pinkham weiterführten und ihre Überlegungen darauf aufbauten (vgl. Kapitel 4 und auch Kapitel 5).

1.2. Historischer Hintergrund



(a) S. Newcomb (1835-1909)



(b) F. Benford (1883-1948)

Abbildung 1.2: Die Entdecker des *Gesetzes der führenden Ziffern*²

6

N.	L.	0	1	2	3	4	5	6	7	8	9	P. P.
200	30	103	125	146	168	190	211	233	255	276	298	
201		320	341	363	384	406	428	449	471	492	514	22 21
202		535	557	578	600	621	643	664	685	707	728	1 2,1 2,1
203		750	771	792	814	835	856	878	899	920	942	2 4,4 4,2
204		963	984	1006	1027	1048	1069	1091	1112	1133	1154	3 6,6 6,3
205	31	175	197	218	239	260	281	302	323	345	366	4 8,8 8,4
206		387	408	429	450	471	492	513	534	555	576	5 11,0 10,5
207		597	618	639	660	681	702	723	744	765	785	6 13,2 12,6
208		806	827	848	869	890	911	931	952	973	994	7 15,4 14,7
209	32	015	035	056	077	098	118	139	160	181	201	8 17,6 16,8
210		222	243	263	284	305	325	346	366	387	408	9 19,8 18,9
211		428	449	469	490	510	531	552	572	593	613	80
212		634	654	675	695	715	736	756	777	797	818	1 2,0
213		838	858	879	899	919	940	960	980	1001	1021	2 4,0
214	33	041	062	082	102	122	143	163	183	203	224	3 6,0
215		244	264	284	304	324	344	364	384	404	424	4 8,0

Abbildung 1.3: Ausschnitt aus einer Logarithmentafel³

²aus <http://www.physics.csbsju.edu/astro/newcomb/SNewcomb.html> (08.01.2016) und

aus <http://www.s9.com/Biography/benford-frank/> (08.01.2016)

³aus <http://zahlwort.blogger.de/stories/515798/> (08.01.2016)

1.3 Newcombs Mantissengesetz

Die stärker abgenutzten vorderen Seiten seiner Logarithmentafeln führten bei Newcomb zu der Erkenntnis, dass die Ziffer 1 als *führende Ziffer* häufiger vorkommen muss, als die restlichen Ziffern und, dass die Häufigkeit zur Ziffer 9 hin abnimmt (vgl. [New81, S. 39]).

Seine Überlegungen dazu, fasste er in seinen *Mantissengesetz* zusammen:

„The law of probability of the occurrence of numbers is such that all mantissae of their logarithms are equally probable.“⁴

Theorem 1.3.1 (Mantissengesetz (vgl. [New81, S. 39]))

Die Häufigkeit von Zahlen ist so, dass die Mantissen ihrer Logarithmen gleichverteilt sind.

Newcomb versteht dabei unter der *Mantisse* einer Zahl ihren *fraktionalen Teil*. (Im Allgemeinen wird der Begriff der Mantisse jedoch wie in Kapitel 5 gebraucht.)

Definition 1.3.2 (Mantisse i.S.v. Newcomb (vgl. [Hü07, S. 3]))

Für $x \in \mathbb{R}^+$ ist die *Mantisse* von x (i.S.v. Newcomb) gegeben durch

$$\langle x \rangle := x - \lfloor x \rfloor \equiv x \pmod{1}.$$

Newcomb führte folgende Überlegungen an, mit denen er zu seinem Mantissengesetz gelangte (vgl. [New81, S. 39 f]).

Jede positive reelle Zahl x lässt sich darstellen als $x := b^d$ mit $d = \log_b x$, wobei b der Basis des betrachteten Zahlensystems entspricht (in der Regel ist b gleich 10). Zwei Zahlen b^{c+s} und $b^{c'+s}$, mit $c, c' \in \mathbb{Z}$ und $s \in \mathbb{R}$, unterscheiden sich nur durch Multiplikation mit einer Potenz von b . Da man nur an den führenden Ziffern interessiert ist, reicht es, Zahlen z zwischen 1 und b zu betrachten. Diese lassen sich durch $z := b^s$ darstellen, wobei s aus $[\log_b(1), \log_b(b)) = [0, 1)$ ist. Die Exponenten s entsprechen dabei den *Mantissen der Logarithmen* der positiven reellen Zahlen x und

⁴[New81, S. 39]

„(...) we may suppose them arranged around a circle according to some law.“⁵

Das Gesetz nach dem diese Exponenten auf dem Kreis verteilt sind, ist nach Newcomb die *Gleichverteilung*.

Einen mathematischen Beweis für seine Überlegungen liefert Newcomb allerdings nicht.

In [Hü07, S. 4] findet sich allerdings eine heuristische Begründung für die Plausibilität des Mantissengesetzes.

Wir betrachten eine Menge von positiven Zufallszahlen (in diesem Fall handelt es sich um Weibull-verteilte Zufallszahlen). Abbildung 1.4 zeigt diese zunächst auf der reellen Zahlengerade abgebildet und dann darauf die Funktion $\log_{10}(\cdot)$ angewendet. Anschließend wird noch die Mantisse dieser Zahlen genommen. Betrachtet man genügend viele Zahlen und deren Logarithmen, so verteilen sich deren Mantissen gleich auf dem Kreis.

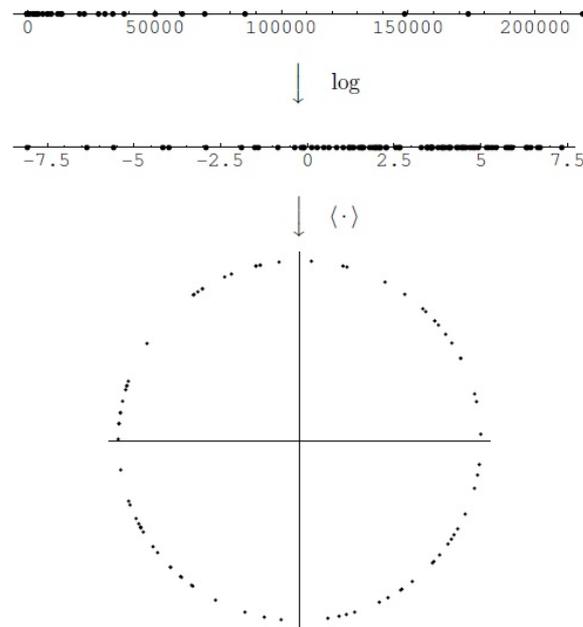


Abbildung 1.4: Heuristische Herleitung des Mantissengesetzes⁶

⁵[New81, S. 40]

⁶aus [Hü07, S. 4]

KAPITEL 1. EINFÜHRUNG

Betrachten wir nun eine Menge von Zufallszahlen, die so verteilt ist, dass sie dem *Mantissengesetz* gehorcht (vgl. im Folgenden [Hü07, S. 3 ff]). Weiter sei

$$E_t := \bigcup_{k \in \mathbb{Z}} 10^k \cdot [t, t+1) \\ = \{x \in \mathbb{R}^+ : \text{führende Ziffer von } x \text{ ist } t\},$$

mit $t \in \{1, 2, \dots, 9\}$.

Zum Beispiel sind die Menge E_1 und die Menge E_2 ausschnittthaft gegeben durch

$$E_1 = \dots \cup [0.1, 0.2) \cup [1, 2) \cup [10, 20) \cup [100, 200) \cup \dots \\ E_2 = \dots \cup [0.2, 0.3) \cup [2, 3) \cup [20, 30) \cup [200, 300) \cup \dots$$

Abbildung 1.5 zeigt die Menge E_1 und die Menge E_2 unter dem Bild der Funktion $\langle \log_{10}(\cdot) \rangle$. Die Funktion $\langle \log_{10}(\cdot) \rangle$ ist also invariant unter Multiplikation des Arguments mit einer Potenz von 10.

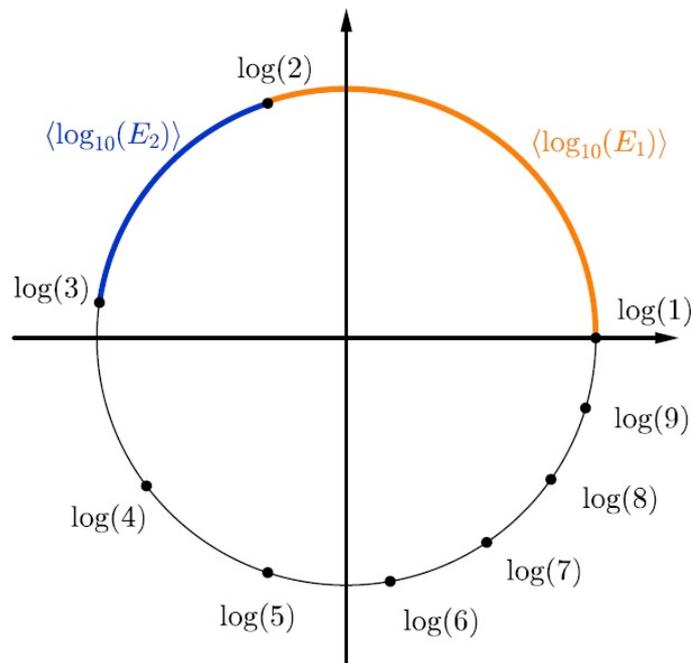


Abbildung 1.5: Menge E_1 und E_2 unter dem Bild der Funktion $\langle \log_{10}(\cdot) \rangle$ ⁷

⁷nach [Hü07, S. 5]

Die Wahrscheinlichkeit, dass nun eine Zufallszahl X mit Ziffer t beginnt, also in der Menge E_t liegt, ist gegeben durch

$$\begin{aligned}
 P(X \in E_t) &= P\left(X \in \bigcup_{k \in \mathbb{Z}} 10^k \cdot [t, t+1)\right) \\
 &= P(\langle \log_{10}(X) \rangle \in [\log_{10} t, \log_{10}(t+1))) \\
 &\stackrel{1.3.1}{=} \log_{10}(t+1) - \log_{10}(t) \\
 &= \log_{10}\left(1 + \frac{1}{t}\right), \tag{1.4}
 \end{aligned}$$

wobei die dritte Zeile aus dem *Mantissengesetz* folgt, wonach $\langle \log_{10}(X) \rangle$ gleichverteilt auf $[0, 1)$ ist.

Aus Newcombs *Mantissengesetz* kann somit *Benfords Gesetz der führenden Ziffer* (vgl. Theorem 1.1.1) gefolgert werden.

1.4 Benfords Erkenntnisse

1.4.1 Empirische Befunde

Als Benford dieselbe Entdeckung mit Logarithmentafeln 57 Jahre später als Newcomb machte, führte er eine große Menge an empirischen Untersuchungen durch, um mehr über dieses Phänomen herauszufinden und die Gesetzmäßigkeit der Verteilung der ersten Ziffern in verschiedensten Anwendungsbereichen zu bestätigen (vgl. im Folgenden [Ben38]).

Er sammelte dabei Daten aus sehr unterschiedlichen Bereichen. So untersuchte er beispielsweise die Oberflächen vieler Seen, Einträge auf Titelseiten von Zeitungen, Adressen, aber auch mathematische Tabellen von Molekular-Gewichten, physikalischen Konstanten und Atomgewichten (vgl. [Ben38, S. 552 ff]). Abbildung 1.6 zeigt die gesammelten Daten Benfords.

Die Spalte *Title* gibt dabei an, welche Daten untersucht wurden. Die gewählten Daten unterscheiden sich stark im Grad ihrer zufälligen Zusammengehörigkeit. Am Ende jeder Spalte findet sich die durchschnittliche Häufigkeit für Zahlen mit der jeweiligen ersten signifikanten Ziffer (in Prozent). Die letzte Spalte zeigt die Anzahl der gesammelten Daten.

PERCENTAGE OF TIMES THE NATURAL NUMBERS 1 TO 9 ARE USED AS FIRST DIGITS IN NUMBERS, AS DETERMINED BY 20,229 OBSERVATIONS

Group	Title	First Digit									Count
		1	2	3	4	5	6	7	8	9	
A	Rivers, Area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Spec. Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
K	n^{-1}, \sqrt{n}, \dots	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8.0	8.9	5000
L	Design	26.8	14.8	14.3	7.5	8.3	8.4	7.0	7.3	5.6	560
M	<i>Digest</i>	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
N	Cost Data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
O	X-Ray Volts	27.9	17.5	14.4	9.0	8.1	7.4	5.1	5.8	4.8	707
P	Am. League	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3.0	1458
Q	Black Body	31.0	17.3	14.1	8.7	6.6	7.0	5.2	4.7	5.4	1165
R	Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	342
S	$n!, n^2 \dots n!$	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
T	Death Rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Average		30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
Probable Error		± 0.8	± 0.4	± 0.4	± 0.3	± 0.2	± 0.2	± 0.2	± 0.2	± 0.3	—

Abbildung 1.6: Benfords gesammelte Daten⁸

Die durchschnittliche Häufigkeit von führenden 1en liegt dabei bei 0,306, was ungefähr dem Logarithmus von 2 entspricht. Die Häufigkeit von führenden 2en liegt bei 0,185, was ein wenig größer ist als der Logarithmus von $\frac{3}{2}$. Ähnliche Beziehungen zwischen den Häufigkeiten und Logarithmen lassen sich für alle führenden Ziffern finden, sodass Benford daraus die bekannte Formel

$$F_a = \log\left(\frac{a+1}{a}\right) \tag{1.5}$$

für die Häufigkeit F_a der führenden Ziffer a festmachte (vgl. [Ben38, S. 554]). Anzumerken ist noch, dass sich die Daten aus Abbildung 1.6 im Allgemeinen aus Zahlen mit mindestens vier Stellen zusammensetzen, denn wie in Abschnitt 1.4.2 gezeigt wird, ist (1.5) ein Verteilungsgesetz für große Zahlen.

⁸aus [Ben38, S. 553]

Betrachtet man die Häufigkeitsverteilungen in Abbildung 1.6 genauer, so stellt man fest, dass Datensätze, deren enthaltene Daten nur wenig zusammenhängen, besser dem Benfordschen Gesetz (1.5) genügen, als Datensätze mit formaler oder mathematischer Struktur. Die beste Übereinstimmung mit dem *Logarithmischen Gesetz* fand Benford bei der Untersuchung von (nicht-ausgeschriebenen) arabischen Zahlen auf Titelseiten von Zeitungen (vgl. Abbildung 1.6, Buchstabe D). Zahlen, die ein Datum beschreiben, wurden dabei nicht berücksichtigt, da diese von vornherein auf einen bestimmten Bereich beschränkt sind. Und da man ausgeschriebene Zahlen nicht berücksichtigte, wurde sichergestellt, dass nur Zahlen größer gleich 10 in den Datensatz aufgenommen wurden. Auch die ersten 342 Straßen Adressen in der (damals) aktuellen Ausgabe des *American Men of Science* zeigen eine sehr gute Übereinstimmung (vgl. Abbildung 1.6, Buchstabe R). Im Gegensatz dazu gibt es größere Abweichungen vom *Logarithmischen Gesetz* bei mathematischen Tabellen zum Atomgewicht (vgl. Abbildung 1.6, Buchstabe J). Abbildung 1.7 zeigt die Summe der Abweichungen zwischen den jeweiligen beobachteten und den berechneten Häufigkeiten in Benfords Datensätzen.

SUMMATION OF DIFFERENCES BETWEEN OBSERVED AND THEORETICAL FREQUENCIES

Order	Item	Nature	Difference	Order	Item	Nature	Difference
1	D	Newspaper Items	2.8	11	N	Cost Data, Concrete	12.4
2	F	Pressure Lost, Air Flow	3.2	12	S	$n^1 \cdots n^8, n!$	13.8
3	G	H.P. Lost in Air Flow	4.8	13	L	Design Data Generators	16.6
4	R	Street Addresses, A.M.S.	5.4	14	B	Population, U. S. A.	16.6
5	P	Am. League, 1936	6.6	15	I	Drainage Rate of Rivers	21.6
6	Q	Black Body Radiation	7.2	16	K	$n^{-1}, \sqrt{n} \cdots$	22.8
7	O	X-Ray Voltage	7.4	17	H	Molecular Wgts.	23.2
8	M	<i>Readers' Digest</i>	8.4	18	E	Specific Heats	24.2
9	A	Area Rivers	9.8	19	C	Physical Constants	34.9
10	T	Death Rates	11.2	20	J	Atomic Weights	35.4

Abbildung 1.7: Summe der Abweichungen zwischen beobachteten und berechneten Häufigkeiten⁹

Betrachten wir exemplarisch die Berechnung der Abweichung zwischen beobachteter und berechneter Häufigkeit für Item D (vgl. Tabelle 1.2).

⁹aus [Ben38, S. 557]

	1	2	3	4	5	6	7	8	9
Item D (in %)	30,0	18,0	12,0	10,0	8,0	6,0	6,0	5,0	5,0
berechn. Wahrsch. (in %)	30,1	17,6	12,5	9,7	7,9	6,7	5,8	5,1	4,6
Abweichung	0,1	0,4	0,5	0,3	0,1	0,7	0,2	0,1	0,4

Tabelle 1.2: Exemplarische Berechnung der Abweichungen zwischen beobachteten und berechneten Häufigkeiten für Item D

Summiert man schließlich die Differenzen aus Tabelle 1.2 auf, so erhält man

$$0,1 + 0,4 + 0,5 + 0,3 + 0,1 + 0,7 + 0,2 + 0,1 + 0,4 = 2,8$$

Die Summe der Abweichungen zwischen beobachteten und berechneten Häufigkeiten für die anderen Items erhält man analog.

1.4.2 Ordnung der natürlichen Zahlen

Benford versuchte natürlich, diese beobachtete logarithmische Verteilung der ersten Ziffern zu erklären.

„There must be some underlying causes that distort what we call the ‘natural’ number system into a logarithmic distribution [...]“¹⁰

Zur Untersuchung dieser Gründe betrachtete Benford zunächst die Häufigkeitsverteilung der natürlichen Zahlen an sich (vgl. [Ben38, S. 557 ff]).

Sehr anschaulich wird diese Häufigkeitsverteilung in [Hum00, S. 139 ff] dargestellt.

Notation 1.4.1 (vgl. [Hum00, S. 139])

$P_n(p)$ bezeichne die Wahrscheinlichkeit, mit der eine Ziffer p als Anfangsziffer unter den ersten n natürlichen Zahlen auftritt. Die Folge $(P_n(p))_{n \in \mathbb{N}}$ repräsentiert somit die Änderungen in der Auftrittswahrscheinlichkeit einer Anfangsziffer p über die natürlichen Zahlen hinweg.

¹⁰[Ben38, S. 557]

Die eben bezeichnete Wahrscheinlichkeit hängt offensichtlich von der „Grundgesamtheit“ n ab.

Zur Untersuchung der Auftretswahrscheinlichkeiten der Ziffer 1 als Anfangsziffer unter den ersten n natürlichen Zahlen betrachten wir die Folge $(P_n(1))_{n \in \mathbb{N}}$.

Für $n = 1$ ist $P_1(1) = 1$, für $n = 2$ ist $P_2(1) = \frac{1}{2}$, usw. Die Wahrscheinlichkeit nimmt ab bis $P_9(1) = \frac{1}{9}$ und steigt dann wieder bis $n = 19$ auf $P_{19}(1) = \frac{11}{19}$. Anschließend fällt sie erneut bis $n = 99$ auf $P_{99}(1) = \frac{1}{9}$ und steigt schließlich wieder auf $P_{199}(1) = \frac{111}{199}$. Dies kann von Zehnerpotenz zu Zehnerpotenz immer so weiter geführt werden. Für wachsendes n pendelt sich die Wahrscheinlichkeit bei *keinem* stabilen Wert ein, nur die Phasen, in denen die Wahrscheinlichkeit abnimmt oder ansteigt, werden länger. Die Folge $(P_n(1))_{n \in \mathbb{N}}$ divergiert also.

$P_n(1)$ schwankt zwischen der gleichbleibenden *Untergrenze* $U_n(1) = \frac{1}{9}$ und einer nur fast gleichbleibenden *Obergrenze* $O_m(1) = \frac{11}{19}, \frac{111}{199}, \frac{1111}{1999}, \dots$, wobei m der Anzahl der Ziffern in Zähler bzw. Nenner entspricht. $O_m(1)$ ist monoton fallend und nach unten beschränkt.

Eine geschlossene Form für $O_m(1)$ ist durch

$$O_m(1) = \frac{5}{9} \cdot \frac{10^m - 1}{10^m - 5}$$

gegeben, anhand der leicht der Grenzwert $\frac{5}{9} = 0.\overline{5}$ abgelesen werden kann.

Abbildung 1.8 zeigt den Verlauf von $P_n(1)$ als kontinuierliche Funktion in Abhängigkeit von n . Die x -Achse (bzw. n -Achse) wurde dabei logarithmisch skaliert, damit die Abstände zwischen den Zehnerpotenzen gleich bleiben.

Betrachten wir nun die Ober- und Untergrenzen der Ziffern $2, 3, \dots, 9$ und greifen dabei noch einmal explizit die Ziffer 9 heraus, d.h. wir betrachten die Folge $(P_n(9))_{n \in \mathbb{N}}$. Für $n = 1, \dots, 8$ ist $P_n(9) = 0$ und für $n = 9$ ist $P_n(9) = \frac{1}{9}$. Die Wahrscheinlichkeit sinkt dann ab bis $n = 89$ auf $P_{89}(9) = \frac{1}{89}$ und steigt anschließend wieder an auf $P_{99}(9) = \frac{1}{9}$ für $n = 99$, usw. $P_n(9)$ schwankt somit zwischen einer festen *Obergrenze* $O_n(9) = \frac{1}{9}$ und der nur fast gleichbleibenden *Untergrenze* $U_m(9) = 0, \frac{1}{89}, \frac{11}{899}, \dots$, wobei m der Anzahl der Einsen im Zähler entspricht. Die Folge $U_m(9)$ ist monoton wachsend und nach oben beschränkt und kann durch

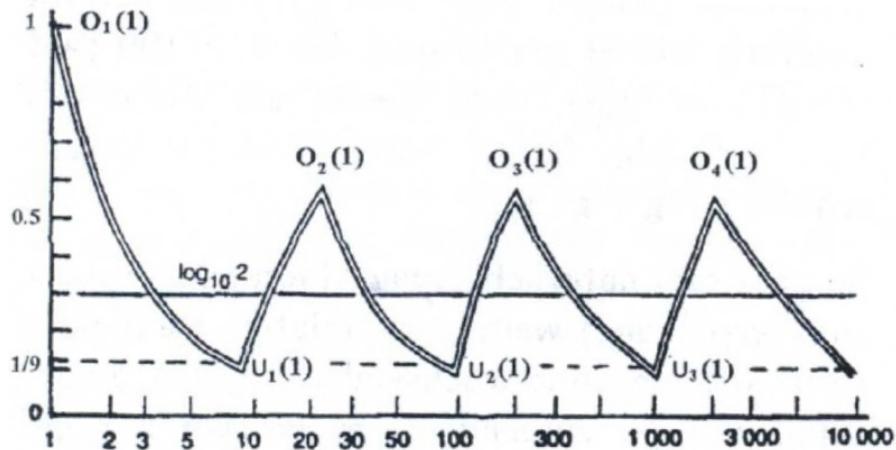


Abbildung 1.8: Grafische Darstellung des Verlaufs von $P_n(1)$ als kontinuierliche Funktion von n ¹¹

die geschlossene Form

$$U_m(9) = \frac{1}{81} \cdot \frac{10^m - 1}{10^m - \frac{1}{9}}$$

dargestellt werden. Sie konvergiert offensichtlich gegen den Wert $\frac{1}{81}$.

Für die Ziffern $p = 2, \dots, 8$ sind weder die Werte der Obergrenze $O_m(p)$ noch der Untergrenze $U_m(p)$ konstant. Allgemein können $U_m(p)$ und $O_m(p)$ für $p = 1, \dots, 9$ durch

$$U_m(p) = \frac{\sum_{k=0}^{m-1} 10^k}{(p-1) \cdot 10^m + \sum_{k=0}^{m-1} 9 \cdot 10^k} = \frac{1}{9p} \cdot \frac{10^m - 1}{10^m - \frac{1}{p}} \quad \text{und}$$

$$O_m(p) = \frac{\sum_{k=0}^{m-1} 10^k}{p \cdot 10^m + \sum_{k=0}^{m-1} 9 \cdot 10^k} = \frac{10}{9(p+1)} \cdot \frac{10^m - 1}{10^m - \frac{10}{p+1}}$$

angegeben werden(, wobei m der Anzahl der Einsen im Zähler bzw. der Anzahl der Neunen im Nenner entspricht).

Abbildung 1.9 zeigt die Verläufe von $P_n(1)$ und $P_n(9)$ als kontinuierliche Funktionen von n . Dabei lässt sich erkennen, dass die 1 als führende Ziffer sehr viel häufiger auftritt als die Ziffer 9.

¹¹aus [Hum00, S. 140]

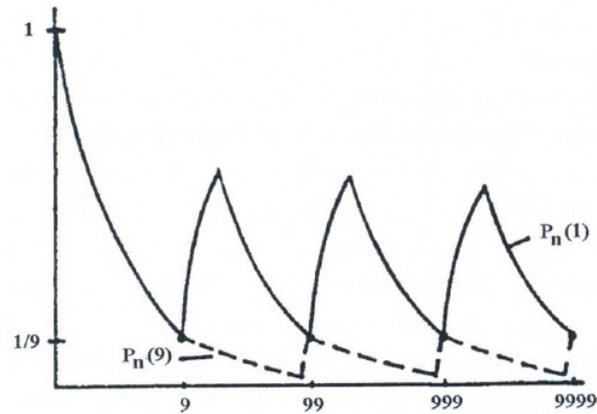


Abbildung 1.9: Gleichzeitige graphische Darstellung der Verläufe von $P_n(1)$ und $P_n(9)$ als kontinuierliche Funktionen von n ¹²

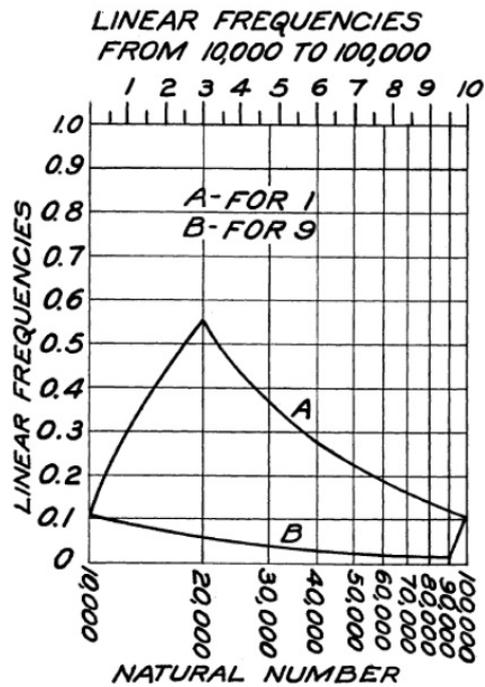


Abbildung 1.10: Häufigkeiten der natürlichen Zahlen zwischen 10.000 und 100.000 ¹³

¹²aus [Hum00, S. 141]

¹³aus [Ben38, S. 559]

Benford selbst zeigt in [Ben38, S. 557 ff] in Abbildung 1.10 einen Ausschnitt der Häufigkeitsverteilung der Anfangsziffer 1 (Kurve *A*) bzw. der Anfangsziffer 9 (Kurve *B*) im Bereich 10.000 und 100.000. Die x -Achse wurde logarithmisch skaliert. Jeder Punkt auf *A* repräsentiert die Häufigkeit von führenden Einsen von 1 bis zum betrachteten Wert n . Berechnet man die Fläche unter der Kurve *A* über den drei Abschnitten 10.000 bis 20.000, 20.000 bis 99.999 und 99.999 bis 100.000, so erhält man sehr genau den Wert 0,3103. (Die Fläche der gesamten Koordinatenebene beträgt 1.) Integration unter der Kurve *A* liefert nach Benford also die *durchschnittliche* Häufigkeit führender Einsen bis zur Zahl 100.000 (vgl. [Ben38, S. 558 f]).

Natürliche Zahlen von 1 bis 10 nannte Benford *first digital order numbers*, Zahlen von 10 bis 100 *second digital order numbers*, usw. Dabei bildet 10 sowohl die letzte Zahl erster Ordnung als auch die erste Zahl zweiter Ordnung (vgl. [Ben38, S. 564 f]).

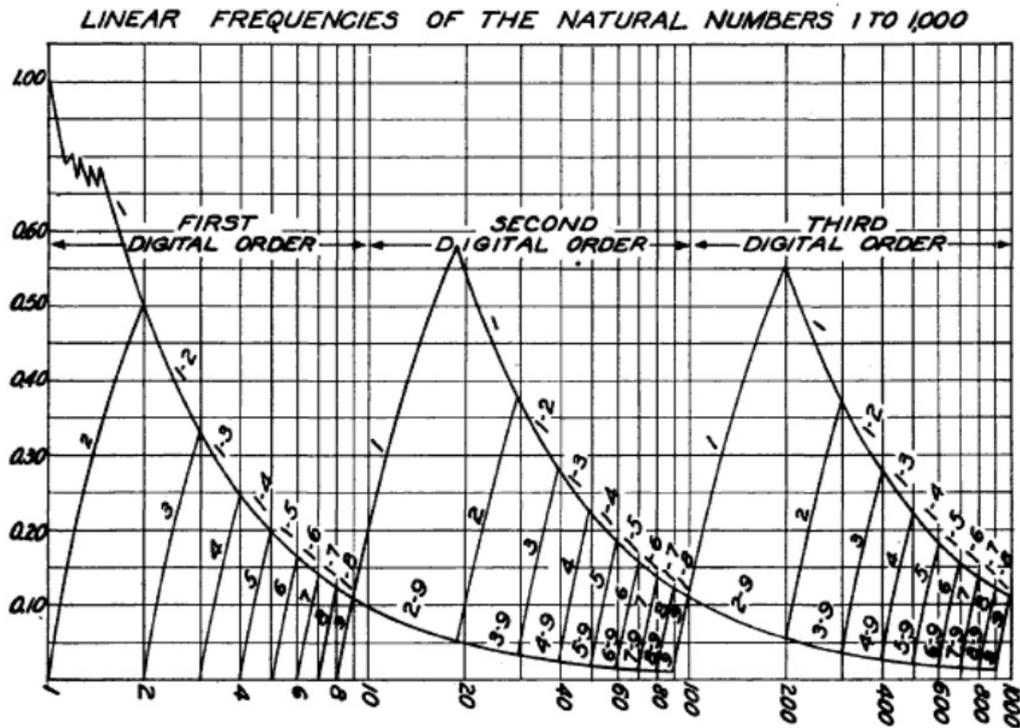
Die Kurven in Abbildung 1.11 zeigen die Häufigkeit, mit der die Anfangsziffern der natürlichen Zahlen von 1 bis n auftreten.

Dabei gehört beispielsweise die Kurve, die von der Zahl 9 aus ansteigt nur zur Ziffer 1, während die von der 9 aus abfallende Kurve für die Ziffern 2 bis 9 steht. Bei 19 steigt die Kurve der Ziffer 2 an und fällt bei 29 mit der Kurve der 1 zusammen. Die Ziffern 1 und 2 haben dann eine gemeinsame Kurve bis die Zahl 99 erreicht ist.

Die betrachteten Kurven werden so gezeichnet, als ob es sich um kontinuierliche Funktionen handelt, obwohl die natürlichen Zahlen eigentlich nur diskrete Werte annehmen. Aber auch diese Tatsache rechtfertigt Benford, indem er sagt:

„[T]he things we use the number system to represent are nearly always perfectly continuous functions, and the number, say 9, given to any phenomenon will be used in some degree for all the infinite sizes of phenomena between 8 and 10 when we confine ourselves to single digit numbers.“¹⁴

¹⁴[Ben38, S. 565]

Abbildung 1.11: Häufigkeiten der natürlichen Zahlen der ersten drei Ordnungen¹⁵

Anhand von Abbildung 1.12 erklärt Benford seine Aussage. Die Linien $h-b$ und $b-j$ zeigen die berechneten Häufigkeiten der Ziffer 1 im Übergangsbereich der Zahlen erster und zweiter Ordnung. Die Häufigkeitskurve der Ziffer 9 (Linie $8-b$), beginnt bereits bei der 8, denn für Zahlen, die minimal größer sind als 8 besteht eine gewisse Wahrscheinlichkeit dafür, eine 9 zu gebrauchen. Bei $8\frac{1}{2}$ ist es sogar gleichwahrscheinlich, diese Zahl 8 oder 9 zu nennen.

Den Flächeninhalt unter der Kurve $8-b-c$ sieht Benford als die Wahrscheinlichkeit, die Zahl 9 für Phänomene in dieser Größenordnung zu verwenden. Nach der Zahl 9 steigt die Kurve $b-j$ der Ziffer 1 an, denn Phänomene zwischen 9 und 10 werden 10 genannt.

¹⁵aus [Ben38, S. 564]

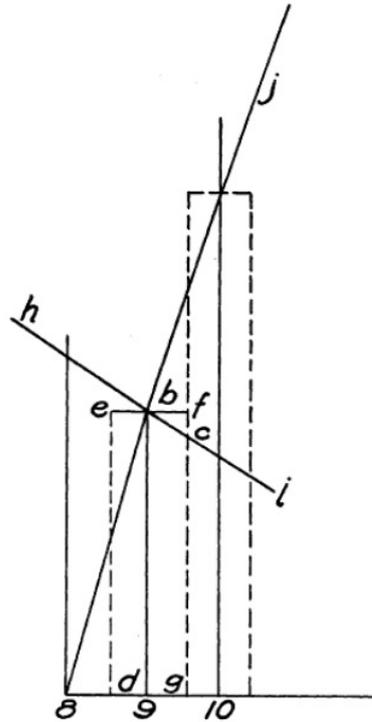


Abbildung 1.12: Vergrößerter Ausschnitt der Häufigkeitskurven im Übergangsbereich zwischen erster und zweiter Ordnung¹⁶

Wie bereits in Abschnitt 1.4.1 angemerkt wurde, ist Benfords *Gesetz der führenden Ziffern* ((1.5), S. 16) ein Verteilungsgesetz für große Zahlen. Für Zahlen höherer Ordnung entsprechen die Flächen unter den Kurven in Abbildung 1.10 den durchschnittlichen Häufigkeiten der ersten Ziffern. Dies verifiziert Benford durch Integration der 1-Kurve im Bereich der Zahlen dritter Ordnung (vgl. Abbildung 1.11) und anschließender Verallgemeinerung auf Zahlen höherer Ordnung (vgl. [Ben38, S. 567 f]).

Die Fläche der 1-Kurve im Bereich der Zahlen *dritter Ordnung* kann als

$$A_1''' = \int_{100}^{199} y_1 dx + \int_{199}^{999} y_2 dx + \int_{999}^{1000} y_3 dx \quad (1.6)$$

geschrieben werden, mit

$$y_1 = \frac{a - 88}{a}, \quad y_2 = \frac{111}{a} \quad \text{und} \quad y_3 = \frac{a - 888}{a}.$$

¹⁶aus [Ben38, S. 565]

Um die Abstände zwischen den Zehnerpotenzen gleich zu halten, wurden die Kurven semilogarithmisch dargestellt mit $x = \ln(a)$, also $dx = \frac{da}{a}$.

Nach Durchführung dieser Substitution und Berechnen der Integrale, erhält man

$$A_1''' = \ln\left(\frac{1990}{999}\right) + \frac{8}{1000}.$$

Analog ergibt sich für die Zahlen der zweiten und ersten Ordnung

$$A_1'' = \ln\left(\frac{190}{99}\right) + \frac{8}{100} \quad \text{und} \quad A_1' = \ln\left(\frac{10}{99}\right) + \frac{8}{10}.$$

Aufgrund der Symmetrie dieser Ergebnisse und den Lösungen für die anderen acht Ziffern formulierte Benford

$$\begin{aligned} F_1^r &= \left[\ln \frac{10(2 \cdot 10^{r-1} - 1)}{10^r - 1} + \frac{8}{10^r} \right] \cdot \frac{1}{\ln 10}, \\ F_{a \neq 1}^r &= \left[\ln \frac{(a+1)10^{r-1} - 1}{a \cdot 10^{r-1} - 1} - \frac{1}{10^r} \right] \cdot \frac{1}{\ln 10} \end{aligned} \tag{1.7}$$

als das *Law of Anomalous Numbers*, wobei r der Ordnung der Zahlen entspricht. Der Faktor $\frac{1}{\ln 10}$ dient dabei nur der Konvertierung des natürlichen in den dekadischen Logarithmus.

Für große Ordnungen r ($r \rightarrow \infty$) kann die $,-1'$ in Zähler und Nenner in (1.7) vernachlässigt werden. Kürzen von 10^r liefert schließlich für $r \rightarrow \infty$

$$\begin{aligned} F_1 &= \log_{10} \frac{2}{1}, \\ F_{a \neq 1} &= \log_{10} \frac{a+1}{a}. \end{aligned} \tag{1.8}$$

Die beiden Gleichungen aus (1.8) können zu einer Gleichung,

$$F_a = \log_{10} \frac{a+1}{a}, \tag{1.9}$$

zusammengefasst werden, was dem von Benford beobachtetem *Gesetz der führenden Ziffern* ((1.5), S. 16) entspricht.

Die Tabelle in Abbildung 1.13 zeigt die berechneten Häufigkeiten für die erste, zweite und dritte Ordnung. Je höher die Ordnung, desto besser stimmen die Werte mit dem Grenzfall ($r \rightarrow \infty$) überein. Die Summe der Häufigkeiten in jeder Spalte liefert den Wert 1.

KAPITEL 1. EINFÜHRUNG

First Digit	First Order	Second Order	Third Order	Limiting Order
	1 to 10	10 to 100	100 to 1000	—
1	0.39319	0.31786	0.30276	0.30103
2	0.25760	0.17930	0.17638	0.17609
3	0.13266	0.12432	0.12487	0.12494
4	0.08152	0.09479	0.09669	0.09691
5	0.05348	0.07631	0.07889	0.07918
6	0.03575	0.06366	0.06662	0.06695
7	0.02352	0.05444	0.05764	0.05799
8	0.01456	0.04742	0.05078	0.05115
9	0.00772	0.04190	0.04537	0.04576

Abbildung 1.13: Berechnete Häufigkeiten der ersten Ziffern in verschiedenen Ordnungen¹⁷

Im Folgenden Abschnitt gehen wir der Frage nach, unter welchen Bedingungen bestimmte Zahlenfolgen dem *Benfordschen Gesetz* genügen.

¹⁷aus [Ben38, S. 569]

KAPITEL 2

Benford Folgen

2.1 Gleichverteilung modulo 1

Folgender Abschnitt orientiert sich an [KN74, S. 1–15].

Definition 2.1.1 (Bruchteil)

Ist $x \in \mathbb{R}$, so wird

$$\langle x \rangle := x - [x] \equiv x \pmod{1}$$

der *Bruchteil* von x genannt. Dabei ist $\langle x \rangle \in [0, 1)$ für alle $x \in \mathbb{R}$.

Bemerkung 2.1.2

Der Bruchteil einer Zahl entspricht der Mantisse einer Zahl i.S.v. Newcomb (vgl. Definition 1.3.2).

Notation 2.1.3

Sei $\omega = (x_n)_{n \in \mathbb{N} \setminus \{0\}}$ eine Folge von reellen Zahlen. Für ein $N \in \mathbb{N} \setminus \{0\}$ und eine Teilmenge $E \subseteq [0, 1)$, sei die Funktion $A(E; N; \omega)$ definiert als die Anzahl der Folgenglieder x_n , $1 \leq n \leq N$, für die $\langle x_n \rangle \in E$ ist.

Wenn klar ist, um welche Folge es sich handelt, wird $A(E; N)$ anstelle von $A(E; N; \omega)$ verwendet.

Definition 2.1.4 (Gleichverteilung modulo 1)

Eine Folge $\omega = (x_n)_{n \in \mathbb{N} \setminus \{0\}}$ von reellen Zahlen ist *gleichverteilt modulo 1*, wenn für alle $a, b \in \mathbb{R}$ mit $0 \leq a < b \leq 1$ gilt

$$\lim_{N \rightarrow \infty} \frac{A([a, b); N; \omega)}{N} = b - a, \quad (2.1)$$

d.h., wenn die relative Anzahl der Folgenglieder deren Bruchteil in einem Intervall $[a, b)$ liegt, gegen die Länge dieses Intervalls konvergiert.

Bemerkung 2.1.5

Sei $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ gleichverteilt modulo 1 (abgekürzt mod 1). Dann kann (2.1) auch geschrieben werden als

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{[a,b)}(\langle x_n \rangle) = \int_0^1 \mathbf{1}_{[a,b)}(\langle x \rangle) dx = b - a \quad (2.2)$$

mit $[a, b) \subseteq [0, 1)$.

Satz 2.1.6

Eine Folge $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ von reellen Zahlen ist gleichverteilt mod 1 genau dann, wenn für jede reellwertige stetige Funktion f , die auf dem abgeschlossenen Intervall $[0, 1]$ definiert ist (d.h. $f \in \mathcal{C}([0, 1])$), gilt

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\langle x_n \rangle) = \int_0^1 f(x) dx. \quad (2.3)$$

Beweis:

1.Fall:

Sei $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ gleichverteilt mod 1. Und sei $f(x) = \sum_{i=0}^{k-1} d_i \mathbf{1}_{[a_i, a_{i+1}]}(x)$ eine Treppenfunktion auf $[0, 1]$ mit $0 = a_0 < a_1 < a_2 < \dots < a_k = 1$. Die Behauptung folgt mit Bemerkung 2.1.5.

2.Fall:

„ \Rightarrow “

Sei $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ gleichverteilt mod 1 und sei $f \in \mathcal{C}([0, 1])$. Sei $\varepsilon > 0$. Dann gibt es nach der Definition des Riemann-Integrals zwei Treppenfunktionen f_1 und f_2 , so dass $f_1(x) \leq f(x) \leq f_2(x)$ für alle $x \in [0, 1]$ mit $\int_0^1 (f_2(x) - f_1(x)) dx \leq \varepsilon$.

Damit gilt

$$\begin{aligned} \int_0^1 f(x) dx - \varepsilon &\leq \int_0^1 f_1(x) dx \\ &\stackrel{1.\text{Fall}}{=} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f_1(\langle x_n \rangle) \leq \liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\langle x_n \rangle) \\ &\leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\langle x_n \rangle) \\ &\leq \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\langle x_n \rangle) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f_2(\langle x_n \rangle) \\ &\stackrel{1.\text{Fall}}{=} \int_0^1 f_2(x) dx \leq \int_0^1 f(x) dx + \varepsilon. \end{aligned}$$

Da ε beliebig war, folgt damit die Behauptung.

„ \Leftarrow “

Sei $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ eine Folge und angenommen, es gelte (2.3) für alle $f \in \mathcal{C}([0, 1])$. Sei $[a, b)$ ein beliebiges Teilintervall von $[0, 1]$. Sei $\varepsilon > 0$. Dann gibt es zwei stetige Funktionen g_1 und g_2 mit $g_1(x) \leq \mathbf{1}_{[a, b)}(x) \leq g_2(x)$ für alle $x \in [0, 1]$ und $\int_0^1 (g_2(x) - g_1(x)) dx \leq \varepsilon$.

Damit gilt

$$\begin{aligned} b - a - \varepsilon &= \int_0^1 \mathbf{1}_{[a, b)}(x) dx - \varepsilon \\ &\leq \int_0^1 g_2(x) dx - \varepsilon \leq \int_0^1 g_1(x) dx \\ &\stackrel{(2.3)}{=} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N g_1(\langle x_n \rangle) \\ &\leq \liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{[a, b)}(\langle x_n \rangle) = \liminf_{N \rightarrow \infty} \frac{A([a, b); N)}{N} \\ &\leq \lim_{N \rightarrow \infty} \frac{A([a, b); N)}{N} \end{aligned}$$

$$\begin{aligned}
 &\leq \limsup_{N \rightarrow \infty} \frac{A([a, b]; N)}{N} \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N g_2(\langle x_n \rangle) \\
 &\stackrel{(2.3)}{=} \int_0^1 g_2(x) dx \leq \int_0^1 g_1(x) dx + \varepsilon \\
 &\leq \int_0^1 \mathbf{1}_{[a, b]}(x) + \varepsilon = b - a + \varepsilon.
 \end{aligned}$$

Da ε beliebig war, folgt mit Definition 2.1.4 die Behauptung. \square

Korollar 2.1.7

Die Folge $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ ist gleichverteilt mod 1 genau dann, wenn jede Riemann-integrierbare Funktion auf $[0, 1]$ die Gleichung (2.3) erfüllt.

Beweis:

„ \Rightarrow “

Klar, nach Satz 2.1.6, da jede Riemann-integrierbare Funktion auf $[0, 1]$ stetig ist.

„ \Leftarrow “

Folgt mit dem 1. Fall des Beweises von Satz 2.1.6. \square

Korollar 2.1.8

Die Folge $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ ist gleichverteilt mod 1 genau dann, wenn für jede komplexwertige auf \mathbb{R} stetige Funktion f mit Periode 1 gilt

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_0^1 f(x) dx. \tag{2.4}$$

Beweis:

„ \Rightarrow “

Wende Satz 2.1.6 auf den Realteil und den Imaginärteil von f an. Die Periodizität von f liefert $f(\langle x_n \rangle) = f(x_n)$.

„ \Leftarrow “

Wähle g_1 und g_2 wie im Beweis von Satz 2.1.6 im 2. Fall so, dass $g_1(0) = g_1(1)$ und $g_2(0) = g_2(1)$ und setze g_1 und g_2 periodisch auf \mathbb{R} fort. Wende im Beweis statt Gleichung (2.3) die Gleichung (2.4) auf g_1 und g_2 an. \square

Satz 2.1.9 (Weyl-Kriterium)

Eine Folge $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ ist *gleichverteilt mod 1* genau dann, wenn

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i h x_n} = 0 \quad \text{für alle } h \in \mathbb{Z} \text{ und } h \neq 0. \quad (2.5)$$

Beweis:

„ \Rightarrow “

Klar, nach Korollar 2.1.8, da $e^{2\pi i h x}$ eine komplexwertige auf \mathbb{R} stetige Funktion mit Periode 1 ist.

„ \Leftarrow “

Sei $(x_n)_{n \in \mathbb{N} \setminus \{0\}}$ eine reelle Folge und es gelte (2.5). Wir zeigen, dass (2.4) für alle komplexwertigen auf \mathbb{R} stetigen Funktionen mit Periode 1 gilt. Mit Korollar 2.1.8 folgt dann, dass x_n gleichverteilt mod 1 ist.

Sei $\varepsilon > 0$ und sei f eine komplexwertige auf \mathbb{R} stetige Funktion mit Periode 1. Mit dem Weierstraß'schen Approximationssatz folgt: Es gibt ein trigonometrisches Polynom $\Psi(x)$ (endliche Linearkombination von Funktionen des Typs $e^{2\pi i h x}$, $h \in \mathbb{Z}$, mit komplexen Koeffizienten), so dass

$$\sup_{0 \leq x \leq 1} |f(x) - \Psi(x)| \leq \varepsilon. \quad (2.6)$$

Es gilt

$$\left| \int_0^1 f(x) dx - \frac{1}{N} \sum_{n=1}^N f(x_n) \right| \leq \left| \int_0^1 (f(x) - \Psi(x)) dx \right| + \left| \int_0^1 \Psi(x) dx - \frac{1}{N} \sum_{n=1}^N \Psi(x_n) \right| + \left| \frac{1}{N} \sum_{n=1}^N (f(x_n) - \Psi(x_n)) \right|.$$

Der erste und der dritte Term auf der rechten Seite sind beide wegen (2.6) $\leq \varepsilon$ unabhängig von der Wahl von N . Wählt man N genügend groß, so ist auch der zweite Term auf der rechten Seite $\leq \varepsilon$ wegen (2.5).

Damit folgt, da ε beliebig war,

$$\int_0^1 f(x) dx = \frac{1}{N} \sum_{n=1}^N f(x_n).$$

Und mit Korollar 2.1.8 folgt nun, dass x_n gleichverteilt mod 1 ist. \square

Beispiel 2.1.10

Sei a irrational. Dann gilt: Die Folge $(na)_{n \in \mathbb{N} \setminus \{0\}}$ ist gleichverteilt mod 1.

Beweis:

Für $h \neq 0$, $h \in \mathbb{Z}$, ist $z := e^{2\pi i h a} \neq 1$ und mit

$$\left| \frac{1}{N} \sum_{n=1}^N e^{2\pi i h n a} \right| \leq \left| \frac{1}{N} \cdot \frac{e^{2\pi i h (N+1)a} - 1}{e^{2\pi i h a} - 1} \right| = \frac{|e^{2\pi i h (N+1)a} - 1|}{N \cdot |e^{2\pi i h a} - 1|} \leq \frac{1}{N \cdot |\sin(\pi h a)|}$$

und Satz 2.1.9 (Weyl-Kriterium) folgt die Behauptung. \square

Satz 2.1.11

Sei $(f(n))_n$, $n = 1, 2, \dots$ eine Folge von reellen Zahlen und sei $f(n)$ gleichverteilt mod 1. Dann gilt

$$\limsup_{n \rightarrow \infty} n |\Delta f(n)| = \infty, \tag{2.7}$$

wobei $\Delta f(n) := f(n+1) - f(n)$ bezeichnet.

Beweis:

Sei $(f(n))_n$, $n = 1, 2, \dots$ gleichverteilt mod 1.

Angenommen

$$\limsup_{n \rightarrow \infty} n |\Delta f(n)| < \infty.$$

Für zwei reelle Zahlen u und v gilt

$$|e^{2\pi i u} - e^{2\pi i v}| \leq 2\pi |u - v|$$

und damit

$$|e^{2\pi i f(n+1)} - e^{2\pi i f(n)}| \leq 2\pi |\Delta f(n)| = \mathcal{O}\left(\frac{1}{n}\right) \quad \text{für } n \rightarrow \infty.$$

Auf der anderen Seite gilt, da $(f(n))_n$ gleichverteilt mod 1 ist nach Satz 2.1.9 (Weyl-Kriterium)

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i h f(n)} = 0.$$

Mit dem *Theorem von Tauber* (vgl. für entsprechende Literatur die Hinweise in [KN74, S. 15]) folgt dann

$$\lim_{N \rightarrow \infty} e^{2\pi i f(n)} = 0,$$

was offensichtlich ein Widerspruch ist, da $e^{2\pi i f(n)}$ eine periodische Funktion ist. \square

2.2 Die Verteilung von führenden Ziffern

Folgender Abschnitt orientiert sich, wenn nicht anders vermerkt, an [Dia77, S. 72–76].

Notation 2.2.1

Sei $D_1 = \{1, 2, \dots, 9\}$ und $D_i = \{0, 1, \dots, 9\}$ für $i \geq 2$ und sei $S_k = \prod_{i=1}^k D_i$. Für $a > 0$ und $x \in S_k$ erhält $a[x]$ den Wert 1, wenn die j -te Ziffer von links von a der Ziffer x_j für $j = 1, \dots, k$ entspricht, ansonsten ist $a[x]$ gleich 0.

Bemerkung 2.2.2

Sei $x \in S_k$. Es gilt

$$\begin{aligned} a[x] = 1 &\iff x_1 10^m + x_2 10^{m-1} + \dots + x_k 10^{m-(k-1)} \\ &\leq a < x_1 10^m + x_2 10^{m-1} + \dots + (x_k + 1) 10^{m-(k-1)} \end{aligned}$$

für ein $m \in \mathbb{N}$, $m > 0$.

Beispiel 2.2.3

Sei $x = (1, 2) \in S_2$. Dann nimmt für $a = 1234$ der Ausdruck $a[(1, 2)]$ den Wert 1 an, da $1200 \leq 1234 < 1300$. Aber für $a = 1432$ ist $a[(1, 2)] = 0$, da $1432 > 1300$ ist.

Definition 2.2.4

Eine Folge von positiven reellen Zahlen $(a_i)_i$, $i = 1, 2, \dots$, ist eine (*starke*) *Benford-*

Folge, wenn

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_i[x] = l(x)$$

für alle $x \in S_k$ und alle $k \in \mathbb{Z}$ mit

$$\begin{aligned} l(x) &= \log_{10} \left(\sum_{i=1}^{k-1} \frac{x_i}{10^{i-1}} + \frac{x_k + 1}{10^{k-1}} \right) - \log_{10} \left(\sum_{i=1}^k \frac{x_i}{10^{i-1}} \right) \\ &= \log_{10} \left(1 + \frac{1}{\sum_{i=1}^k x_i \cdot 10^{k-i}} \right), \end{aligned}$$

d.h. die relative Anzahl an Folgengliedern, die mit der Ziffernfolge $x_1 \dots x_k$ beginnen, beträgt $l(x)$ und entspricht somit der Benford-Wahrscheinlichkeit.

Beispiel 2.2.5

Für $k = 1$ entspricht die relative Häufigkeit der führenden Ziffer x_1 in einer (starken) Benford-Folge

$$\log_{10} \left(1 + \frac{1}{x_1} \right).$$

Satz 2.2.6

Die Folge $(a_i)_i$, $i = 1, 2, \dots$, mit $a_i > 0$, ist eine (starke) Benford-Folge genau dann, wenn die Folge $(\log_{10} a_i)_i$ gleichverteilt mod 1 ist.

Beweis:

„ \Leftarrow “

Sei $k \in \mathbb{N}$ und sei $x \in S_k$. Dann gilt für ein $j \in \mathbb{N}$, $j > 0$

$$\begin{aligned} a_i[x] = 1 &\iff x_1 10^j + x_2 10^{j-1} + \dots + x_k 10^{j-(k-1)} \\ &\leq a_i < x_1 10^j + x_2 10^{j-1} + \dots + (x_k + 1) 10^{j-(k-1)} \\ &\iff x_1 + \frac{x_2}{10} + \frac{x_3}{10^2} + \dots + \frac{x_k}{10^{k-1}} \\ &\leq \frac{a_i}{10^j} < x_1 + \frac{x_2}{10} + \frac{x_3}{10^2} + \dots + \frac{x_k + 1}{10^{k-1}} \\ &\iff \log_{10} \left(x_1 + \frac{x_2}{10} + \frac{x_3}{10^2} + \dots + \frac{x_k}{10^{k-1}} \right) \end{aligned}$$

$$\begin{aligned}
 &\leq \underbrace{\log_{10} \left(\frac{a_i}{10^j} \right)} < \log_{10} \left(x_1 + \frac{x_2}{10} + \frac{x_3}{10^2} + \dots + \frac{x_k + 1}{10^{k-1}} \right) \\
 &= \log_{10}(a_i) - \log_{10} 10^j \\
 &= \log_{10}(a_i) - j \\
 &= \langle \log_{10}(a_i) \rangle \\
 &\iff \log_{10} \left(x_1 + \dots + \frac{x_k}{10^{k-1}} \right) \leq \langle \log_{10} a_i \rangle < \log_{10} \left(x_1 + \dots + \frac{x_k + 1}{10^{k-1}} \right)
 \end{aligned}$$

Da $(\log_{10} a_i)_i$ (nach Annahme) gleichverteilt mod 1 ist, folgt mit Bemerkung 2.1.5

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[a,b)}(\langle \log_{10} a_i \rangle) = b - a \quad (2.8)$$

$$\text{mit } [a, b) = \left[\log_{10} \left(\sum_{i=1}^{k-1} \frac{x_i}{10^{i-1}} + \frac{x_k + 1}{10^{k-1}} \right), \log_{10} \left(\sum_{i=1}^k \frac{x_i}{10^{i-1}} \right) \right).$$

Bleibt noch zu zeigen:

$$\mathbf{1}_{[a,b)}(\langle \log_{10} a_i \rangle) = a_i[x]. \quad (2.9)$$

Es gilt

$$\begin{aligned}
 \mathbf{1}_{[a,b)}(\langle \log_{10} a_i \rangle) &\iff \mathbf{1}_{[a,b)}(\log_{10} a_i - j) \quad \text{für ein } j \in \mathbb{Z} \\
 &\iff \mathbf{1}_{[a,b)}(\log_{10} a_i - \log_{10} 10^j) \\
 &\iff \mathbf{1}_{[a,b)} \left(\log_{10} \frac{a_i}{10^j} \right) \\
 &\iff \mathbf{1}_{\left[x_1 + \dots + \frac{x_k}{10^{k-1}}, x_1 + \dots + \frac{x_k + 1}{10^{k-1}} \right)} \left(\frac{a_i}{10^j} \right) \\
 &\iff \mathbf{1}_{\left[x_1 10^j + \dots + x_k 10^{j-(k-1)}, x_1 10^j + \dots + (x_k + 1) 10^{j-(k-1)} \right)}(a_i) = a_i[x].
 \end{aligned}$$

Mit (2.8) und (2.9) folgt schließlich

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_i(x) = l(x).$$

„ \Rightarrow “

Jedes Intervall $[a, b)$, $0 \leq a < b \leq 1$, kann beliebig genau durch eine endliche Vereinigung von Intervallen mit Endpunkten der Form $\log_{10}\left(x_1 + \frac{x_2}{10} + \dots + \frac{x_k}{10^{k-1}}\right)$ angenähert werden.

Wenn (a_i) eine (starke) Benford-Folge ist, dann gilt nach Definition 2.2.4

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_i(x) = l(x) \quad (\text{mit } l(x) \text{ wie in der Definition})$$

und mit (2.9) folgt (2.8). D.h. die relative Anzahl an Folgengliedern $(\log_{10} a_i)$, die in ein Intervall mit solchen Endpunkten fällt, ist proportional zur Länge des Intervalls.

Damit muss die Folge $(\log_{10} a_i)$ gleichverteilt mod 1 sein. \square

Bemerkung 2.2.7 (vgl. [BH15, S. 44])

Ist die Folge $(a_i)_i$, $i = 1, 2, \dots$ eine Folge von reellen Zahlen, enthält also u.U. auch negative Folgenglieder, dann betrachte die Folge $(\log_{10} |a_i|)_i$.

Beispiel 2.2.8

Die Folge $a_n := q^n$, mit $\xi := \log_{10} q$ irrational, ist eine (starke) Benford-Folge.

Beweis:

Die Folge $\log_{10} a_n = n\xi$ ist gleichverteilt mod 1 nach Beispiel 2.1.10. Mit Satz 2.2.6 folgt dann die Behauptung. \square

Korollar 2.2.9

Wenn eine Folge (a_i) , $i = 1, 2, \dots$, von positiven reellen Zahlen eine (starke) Benford-Folge ist, dann gilt

$$\limsup_{n \rightarrow \infty} \left(n \log_{10} \left(\frac{a_{n+1}}{a_n} \right) \right) = \infty.$$

Beweis:

Da (a_i) eine (starke) Benford-Folge ist, ist nach Satz 2.2.6 die Folge $(\log_{10} a_i)$ gleichverteilt mod 1.

Die Behauptung folgt dann mit Satz 2.1.11. \square

Dieses Korollar ist ein nützliches Hilfsmittel, um zu verifizieren, dass eine gegebene Folge *keine* Benford-Folge ist.

Beispiel 2.2.10 (vgl. [BGM11, S. 45])

- a) Die Folge $a_n := n^b$, mit $b \in \mathbb{R}$, ist *keine* (starke) Benford-Folge.
- b) Die Folge $b_n := \log_b n$, mit $b \in \mathbb{R}$, $b > 1$, ist *keine* (starke) Benford-Folge.

Beweis:

zu a): Es gilt

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left(n \log_{10} \left(\frac{a_{n+1}}{a_n} \right) \right) &= \limsup_{n \rightarrow \infty} \left(n \log_{10} \left(\frac{(n+1)^b}{n^b} \right) \right) \\ &= \limsup_{n \rightarrow \infty} \left(n (\log_{10}(n+1)^b - \log_{10} n^b) \right) \\ &= \limsup_{n \rightarrow \infty} (b \cdot n \log_{10}(n+1) - b \cdot n \log_{10} n) = 0, \end{aligned}$$

da $\log_{10}(n+1)$ und $\log_{10} n$ für genügend großes n „fast denselben Wert annehmen“.

zu b): Es gilt

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left(n \log_{10} \left(\frac{a_{n+1}}{a_n} \right) \right) &= \limsup_{n \rightarrow \infty} \left(n \log_{10} \left(\frac{\log_b(n+1)}{\log_b n} \right) \right) \\ &= \limsup_{n \rightarrow \infty} (n \log_{10}(\log_b(n+1)) - n \log_{10}(\log_b(n))) \neq \infty. \end{aligned}$$

□

Ein weiteres Beispiel für eine Benford-Folge ist die Folge der *Fibonacci-Zahlen*. Um dies zu zeigen, ist noch ein zusätzlicher Satz notwendig.

Satz 2.2.11 (vgl. [BH15, S. 53])

Sei $(b_n)_n$, $n = 1, 2, \dots$, eine Folge von reellen Zahlen, so dass $\lim_{n \rightarrow \infty} \left| \frac{b_n}{a^n} \right|$ existiert und positiv ist für ein $a > 0$. Dann ist $(b_n)_n$ eine *Benford-Folge* genau dann, wenn $\log_{10} |a|$ irrational ist.

Beweis:

Zeige:

$$\begin{aligned} (b_n) \text{ ist eine Benford-Folge} &\stackrel{1)}{\iff} (a^n) \text{ ist eine Benford-Folge} \\ &\stackrel{2)}{\iff} (n \log_{10} a) \text{ ist gleichverteilt mod } 1 \\ &\stackrel{3)}{\iff} (\log_{10} a) \text{ ist irrational.} \end{aligned}$$

zu 1): Folgt mit Satz 4.12(i) in [BH15, S. 51 f].

zu 2): Folgt aus Satz 2.2.6.

zu 3):

„ \Leftarrow “ Folgt aus Beispiel 2.1.10.

„ \Rightarrow “ Nach Satz 2.1.9 (Weyl-Kriterium) gilt, da $(\log_{10} a)$ nach Voraussetzung gleichverteilt mod 1 ist:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i h (n \log_{10} a)} = 0 \quad \text{für alle } h \in \mathbb{Z}, h \neq 0. \quad (2.10)$$

Angenommen $c := \log_{10} a$ ist rational, also $c = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N} \setminus \{0\}$.

Dann ist für $h = p$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \underbrace{e^{2\pi i p n}}_{=1} = 1 \neq 0,$$

was ein Widerspruch zu (2.10) ist. Damit muss $\log_{10} a$ irrational sein. \square

Beispiel 2.2.12 (vgl. [BH15, S. 53 f])

Die Folge der *Fibonacci-Zahlen* $(F_n) = (1, 1, 2, 3, 5, 8, 13, \dots)$ ist eine Benford-Folge.

Beweis:

Die Folge der Fibonacci-Zahlen kann explizit durch die Formel

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1}{2} + \frac{1}{2}\sqrt{5} \right)^n - \left(\frac{1}{2} - \frac{1}{2}\sqrt{5} \right)^n \right) = \frac{\varphi^n - \frac{(-1)^n}{\varphi^n}}{\sqrt{5}},$$

mit $n \in \mathbb{N} \setminus \{0\}$ und $\varphi := \frac{1}{2}(1 + \sqrt{5})$, dargestellt werden.

Es ist $\varphi > 0$ und es gilt

$$\lim_{n \rightarrow \infty} \left| \frac{\varphi^n - \frac{(-1)^n}{\varphi^n}}{\sqrt{5} \cdot \varphi^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{1 - \frac{(-1)^n}{\varphi^{2n}}}{\sqrt{5}} \right| = \frac{1}{\sqrt{5}} > 0.$$

2.2. Die Verteilung von führenden Ziffern

Damit und mit der Tatsache, dass $\log_{10} \varphi$ irrational ist, folgt mit Satz 2.2.11 die Behauptung. \square

Im Folgenden wollen wir uns nicht mehr nur auf Folgen beschränken und untersuchen, ob diese dem *Benfordschen Gesetz* genügen. Von nun an betrachten wir die positiven reellen Zahlen \mathbb{R}^+ und zeigen, dass auch hier das *Gesetz der führenden Ziffern* gilt.

KAPITEL 3

Die Methode von Pinkham

Nach Roger S. Pinkham (vgl. im Folgenden [Pin61, S. 1223–1226]), Mathematik-Professor der Rutgers University in New Brunswick (New Jersey)¹⁸, ändert sich die Zahl der physikalischen Konstanten zwar täglich, die Gesamtheit dieser Konstanten unterliegt jedoch einer bestimmten nicht näher definierten Verteilung.

Betrachtet man die Verteilung der Gesamtheit aller dieser physikalischen Konstanten und die daraus abgeleitete Verteilung der führenden Ziffern, so ist es vernünftig anzunehmen, dass sich die Verteilung der ersten Ziffern nicht ändert, wenn man alle diese Konstanten mit einem Faktor multipliziert, denn Einheiten von physikalischen Konstanten sind ja nicht durch die Natur vorgegeben, sondern vom Menschen willkürlich festgelegt. Daher sollte die Verteilung nicht von den verwendeten Einheiten (z.B. Meter oder Zoll) abhängen. Umwandlungen in eine andere Einheit entsprechen im Allgemeinen einer Multiplikation mit einer positiven Konstanten.

Nach Pinkham reicht die Annahme der *Skaleninvarianz* und die Annahme der *Existenz* einer *Verteilungsfunktion* F aller möglichen physikalischen Konstanten aus, um zu zeigen, dass die Wahrscheinlichkeit D , mit der eine positive Konstan-

¹⁸https://it.wikipedia.org/wiki/Roger_Pinkham (23.02.2016)

te eine Anfangsziffer kleiner gleich $n - 1$, mit $n = 2, \dots, 10$ hat, eindeutig durch $\log_{10}(n + 1)$ gegeben ist.

Sei $F : \mathbb{R}^+ \rightarrow [0, 1]$ die kumulative Verteilungsfunktion der Gesamtheit der nicht-negativen physikalischen Konstanten. Dabei soll gelten:

- (1) $\lim_{x \rightarrow 0^+} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$,
- (2) F ist stetig und monoton wachsend.

Die Funktion F induziert ein σ -additives Wahrscheinlichkeitsmaß Φ auf \mathbb{R}^+ . Für jedes halboffene Intervall $[a, b)$ wird $\Phi([a, b)) = F(b) - F(a)$ als die Wahrscheinlichkeit interpretiert, dass eine Zahl im Intervall $[a, b)$ liegt (vgl. [Rai69, S. 344]).

Im Folgenden bezeichne D die Wahrscheinlichkeit, dass eine Zahl in $D_x := \bigcup_{m=-\infty}^{\infty} [10^m, x \cdot 10^m)$, mit x aus \mathbb{R}^+ , liegt.

Dann ist

$$D(x) = \Phi(D_x) = \sum_{m=-\infty}^{\infty} [F(x10^m) - F(10^m)], \quad x > 0 \quad (3.1)$$

eine für positive x wohldefinierte Funktion und $D(n)$ mit $n = 2, \dots, 10$ gibt die Wahrscheinlichkeit an, dass eine Konstante eine Anfangsziffer $\leq n - 1$ hat, da alle Zahlen zwischen 10^m und $n \times 10^m$ mit Ziffern $\leq (n - 1)$ beginnen.

Nach Pinkham gilt nun für die Mengen D_n die Skaleninvarianz:

- (3) $\Phi(D_n) = \Phi(k \cdot D_n)$, für $n = 2, \dots, 10$ und für alle $k \in \mathbb{R}^+$

Geht man davon aus, dass alle physikalischen Konstanten mit einer positiven Konstante c multipliziert wurden, so ergibt sich mit der vorausgesetzten Skaleninvarianz

$$\begin{aligned} D(n) &= \Phi(D_n) = \Phi\left(\frac{1}{c} \cdot D_n\right) = \sum_{m=-\infty}^{\infty} \left[F\left(\frac{n}{c} 10^m\right) - F\left(\frac{10^m}{c}\right) \right] \\ &= \sum_{m=-\infty}^{\infty} \left[F\left(\frac{n}{c} 10^m\right) - F(10^m) + F(10^m) - F\left(\frac{10^m}{c}\right) \right] \\ &= \sum_{m=-\infty}^{\infty} \left[F\left(\frac{n}{c} 10^m\right) - F(10^m) \right] + \sum_{m=-\infty}^{\infty} \left[-F\left(\frac{10^m}{c}\right) + F(10^m) \right] \\ &= D\left(\frac{n}{c}\right) - D\left(\frac{1}{c}\right). \end{aligned}$$

Also somit

$$D(n) = D\left(\frac{n}{c}\right) - D\left(\frac{1}{c}\right) \quad \text{mit } c > 0 \text{ und } n = 2, \dots, 10. \quad (3.2)$$

Nimmt man an, dass (3.2) für beliebige reelle Zahlen n gilt, so kann man unter der zusätzlichen Annahme von Stetigkeit, nach Pinkham, sofort folgern, dass

$$D(n) = \log_{10} n$$

ist.

Dieser Schluss kann auch bereits unter schwächeren Voraussetzungen, als denen, die für (3.2) verwendet wurden gezeigt werden. Dazu formulierte Pinkham folgenden Satz (vgl. [Pin61, S. 1226]):

Satz 3.0.13

Wenn gilt:

1. $D(2) + D(x) = D(2x), \quad x > 0;$
2. $D(10) + D(x) = D(10x), \quad x > 0;$
3. $D(x)$ ist stetig;
4. $D(10) = 1;$

Dann ist $D(x) = \log_{10}(x), \quad x > 0.$

Anmerkung

Man beachte, dass 1. und 2. lediglich Spezialfälle von (3.2), unter der Annahme der Gültigkeit von (3.2) für beliebige reelle Zahlen, sind.

Hinweis: Im Folgenden wird die Basis 10 bei $\log_{10} x$ weggelassen und stattdessen $\log x$ geschrieben.

Beweis:

Sei $H(x) := D(10^x).$

Ziel des Beweises ist zu zeigen, dass $H(y) = y$ für alle y in \mathbb{R} , denn damit folgt

die Behauptung $D(x) = \log x$ für alle $x \in \mathbb{R}^+$.

Mit 1. und 2. folgt

$$H(\log 2) + H(y) = H(\log 2 + y) \quad \text{und} \quad (3.3)$$

$$H(\log 10) + H(y) = H(\log 10 + y) \quad (3.4)$$

mit $-\infty < y < \infty$, da

$$\begin{aligned} H(\log n) + H(y) &= D(10^{\log n} + D(\underbrace{10^y}_{=:x})) \\ &= D(n) + D(z) \stackrel{1. \text{ bzw. } 2.}{=} D(nx) \\ &= D(n \cdot 10^y) = H(\log(n \cdot 10^y)) \\ &= H(\log n + \log 10^y) \\ &= H(\log n + y) \quad \text{für } n = 2, 10. \end{aligned}$$

Daraus erhält man

$$H(N \log n) = N \cdot H(\log n) \quad \text{für } N \in \mathbb{Z} \text{ und } n = 2, 10, \quad (3.5)$$

da

$$\begin{aligned} H(N \log n) &= H(\log n^N) = H(\log n + \log n^{N-1}) \\ &= H(\log n) + H(\log n^{N-1}) \\ &= H(\log n) + H(\log n) + H(\log n^{N-2}) \\ &= \underbrace{H(\log n) + \dots + H(\log n)}_{N\text{-mal}} \\ &= N \cdot H(\log n). \end{aligned}$$

Ebenfalls kann aus den vorliegenden Regeln abgeleitet werden, dass

$$H(N) = N \quad \text{für } N \in \mathbb{Z}. \quad (3.6)$$

Denn mit $H(1) = 1$ (folgt aus 4.) gilt

$$H(N) = H(\log 10^N) = H(N \log 10) \stackrel{(3.5)}{=} N \cdot \underbrace{H(\log 10)}_{=1} = N \cdot H(1) = N \cdot 1 = N.$$

KAPITEL 3. DIE METHODE VON PINKHAM

Aus der Theorie der Kettenbrüche (vgl. dazu [HW54]) weiß man, dass $\log 2$ mittels k -ten Konvergenten sehr gut approximiert werden kann. Es gibt Folgen $(p_k)_k$ und $(q_k)_k$ in \mathbb{Z} , deren Folgenglieder mittels Rekursion¹⁹ bestimmt werden können. Für $k \rightarrow \infty$ gilt für die Approximation

$$\log 2 = \frac{p_k}{q_k} + o\left(\frac{1}{q_k}\right). \quad (3.7)$$

D.h.

$$q_k \log 2 = p_k + o(1) \quad \text{mit } k \rightarrow \infty \quad (3.8)$$

und

$$q_k H(\log 2) = p_k + o(1), \quad (3.9)$$

da

$$\begin{aligned} q_k H(\log 2) &= H(q_k \log 2) = H(p_k + o(1)) = H(p_k) + o(1) \\ &= D(10^{p_k}) + o(1) \stackrel{2}{=} p_k D(10) + o(1) \stackrel{4}{=} p_k + o(1). \end{aligned}$$

Mit (3.8) und (3.9) folgt schließlich

$$H(\log 2) = \log 2. \quad (3.10)$$

Aus Kapitel 2 Beispiel 2.1.10 weiß man, dass die Folge $a_n := (na)_{n \in \mathbb{N} \setminus \{0\}}$ für alle irrationalen a gleichverteilt mod 1 ist, d.h. $a'_n = na - \lfloor na \rfloor$ ist gleichverteilt auf $[0, 1)$

Daher existiert zu jedem festen h , $0 \leq h < 1$, eine Teilfolge a_k , $k \in \mathbb{N}$, die gegen dieses h konvergiert.

Sei $a := \log 2$. Dann ist $a_k := k \log 2 - \lfloor k \log 2 \rfloor$.

Es gilt

$$\begin{aligned} H(a_k) &= H(k \log 2 - \lfloor k \log 2 \rfloor) = D(10^{k \log 2 - \lfloor k \log 2 \rfloor}) \\ &= D((10^{\log 2})^k \cdot 10^{-\lfloor k \log 2 \rfloor}) = D(2^k \cdot 10^{-\lfloor k \log 2 \rfloor}) \\ &= kD(2) + D(10^{-\lfloor k \log 2 \rfloor}) \\ &= kH(\log 2) + H(-\lfloor k \log 2 \rfloor) \stackrel{(3.6)}{=} kH(\log 2) - \lfloor k \log 2 \rfloor \\ &\stackrel{(3.9)}{=} k \log 2 - \lfloor k \log 2 \rfloor = a_k. \end{aligned}$$

¹⁹siehe dazu: <http://mathworld.wolfram.com/ContinuedFraction.html> (03.03.2016)

Somit ist

$$H(a_k) = a_k. \quad (3.11)$$

Da D (nach 4.) als stetig angenommen wird, so ist auch H stetig und es gilt für $k \rightarrow \infty$

$$H(h) = H(\lim_{k \rightarrow \infty} a_k) = \lim_{k \rightarrow \infty} H(a_k) \stackrel{(3.11)}{=} \lim_{k \rightarrow \infty} a_k = h \quad (3.12)$$

für alle $k \in [0, 1)$.

Schließlich folgt mit $y = \underbrace{[y] + y - [y]}_{\in [0,1)}$ und (3.12)

$$\begin{aligned} H(y) &= D(10^{[y]+y-[y]}) = [y] \cdot D(10) + D(10^{y-[y]}) \\ &= [y] \cdot H(y - [y]) = [y] + y - [y] = y \end{aligned} \quad (3.13)$$

für alle $y \in \mathbb{R}$.

Somit gilt

$$H(y) = D(10^y) = y.$$

Substituiere $y := \log x$, $x \in \mathbb{R}^+$.

Damit erhält man

$$D(10^{\log x}) = D(x) = \log x \quad (3.14)$$

für $x \in \mathbb{R}^+$.

Und damit folgt die Behauptung. \square

Mit der Forderung nach Skaleninvarianz und der Existenz einer Verteilungsfunktion für alle möglichen positiven physikalischen Konstanten bewies Pinkham, dass es nur eine einzige mögliche Verteilungsfunktion für die erste Ziffer geben kann, nämlich

$$D(n) = \log_{10} n \quad \text{mit } n = 2, \dots, 10$$

Diese Verteilungsfunktion gibt also die Wahrscheinlichkeit an, dass die erste Ziffer einer Zahl kleiner gleich $(n - 1)$ ist.

KAPITEL 4

Die Methode von Raimi

4.1 Kritik an Pinkham

Ralph A. Raimi (*1924, Detroit), emeritierter Professor der Universität Rochester²⁰ folgert allerdings, aus der von Pinkham geforderten Skaleninvarianz, dass es eine Verteilungsfunktion F (vgl. Kapitel 3, S. 40) aller physikalischen Konstanten *nicht* geben kann (vgl.[Rai69, S. 343 ff]). Er schreibt dazu:

„Pinkham’s hypothesis of scale-invariance is reasonable, given that there *is* a distribution function F describing the occurrence of members of \mathbb{R}^+ in all the possible physical tables in all possible worlds. But such a function cannot exist.“²¹

Wenn die geforderte Skaleninvarianz in Bezug auf die erste Ziffer gültig ist, d.h. für die Mengen D_n , dann scheint es auch vernünftig, zu fordern, dass diese Invarianzeigenschaft auch für jede beliebige (messbare) Teilmenge $A \subset \mathbb{R}^+$ gelten soll,

²⁰vgl. <http://www.math.rochester.edu/people/faculty/rarm/vitae.html> (23.02.2016)

²¹[Rai69, S. 344]

d.h.

$$P(x \in A) = \Phi(A) = \Phi(k \cdot A) = P(x \in k \cdot A), \quad \text{mit } k \in \mathbb{R}^+. \quad (4.1)$$

(Für die Definitionen von Φ und D_n vergleiche Kapitel 3, S. 41.)

Denn was wäre so besonders an den führenden Ziffern? Allerdings erhält man mit (4.1) unmittelbar für alle $k \geq 0$

$$F(1) = P(x \in (0, 1]) = P(x \in (0, k]) = F(k).$$

Das hieße, dass F konstant ist, was aber ein Widerspruch zu den angenommenen Eigenschaften von F ist (vgl. Kapitel 3, S. 41).

Raimi führt noch einen weiteren Grund an, der gegen die Existenz einer Verteilungsfunktion aller physikalischen Konstanten spricht. Dabei handelt es sich eher um einen „philosophischen Grund [. . .], der auch dann greift, wenn man nicht bereit ist, der angesprochenen Invarianzeigenschaft über die Mengen $[D_n]$ hinaus Gültigkeit zu verleihen.“²²

Würde eine derartige Verteilungsfunktion existieren, so gäbe es auch einen Wert $T > 0$ mit $F(T) = P(x \leq T) = P(x \in (0, T]) = \frac{1}{2}$. „D.h. es gäbe einen Wert T mit der Eigenschaft, daß er genau in der Mitte des Universums aller denkbaren Konstanten aller denkbaren Welten läge.“²³ T wäre also der „Median aller denkbaren Konstanten“.²⁴

Dazu meint allerdings Raimi:

T takes on a mystical significance. I cannot bring myself to believe in such a number [. . .].²⁵

„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.“²⁶

²²[Hum00, S. 146]

²³ebd.

²⁴ebd.

²⁵[Rai69, S. 344]

²⁶[Hum00, S.145]

4.2 Banach-Maß und skaliertes Maß

„Raimi selbst [benutzt] sogenannte *Banach-Maße* und *skalierte Maße*, um die Existenz einer oben angesprochenen Verteilungsfunktion aller möglichen positiven Konstanten und eines daraus resultierenden „Mittelpunktes“ all dieser zu vermeiden.“²⁷

„What I shall present is an „unorthodox“ probability model related to Pinkham’s.“²⁸

Raimi (vgl. [Rai69, S. 344 ff]) verwendet in seinen Überlegungen ein endliches *nicht* σ -additives Wahrscheinlichkeits-Maß. „Sein Ziel war [allerdings] dasselbe, [wie Pinkham] zu beweisen:

Wenn die Wahrscheinlichkeit $P(1. \text{ Ziffer} \leq p)$ *skaleninvariant* sein soll, so kommt dafür nur eine logarithmische Verteilungsfunktion in Frage: $P(1. \text{ Ziffer} \leq p) = \log(p + 1)$.“²⁹

Zunächst werden ein paar grundlegende Definitionen gegeben, bevor dann das *Banach-Maß* und das *skalierte Maß* eingeführt werden.

4.2.1 Grundlagen

Die folgenden Definitionen und Sätze stammen aus [Tap13].

Sei Ω eine beliebige *Menge* und bezeichne $\mathcal{P}(\Omega)$ die *Potenzmenge* von Ω . Eine Teilmenge $\mathcal{A} \subset \mathcal{P}(\Omega)$ der Potenzmenge wird auch als *Mengensystem* bezeichnet.

Definition 4.2.1 (σ -Algebra)

Ein Mengensystem $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt eine σ -Algebra, falls gilt:

- i) $\Omega \in \mathcal{A}$.
- ii) Für jede Teilmenge $A \in \mathcal{A}$ gilt: $A^c \in \mathcal{A}$.
- iii) Für jede Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ gilt: $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

²⁷[Hum00, S. 147]

²⁸[Rai69, S. 344]

²⁹[Hum00, S. 147]

Beispiel 4.2.2

Die Potenzmenge $\mathcal{P}(\Omega)$ ist eine σ -Algebra über Ω .

Definition 4.2.3 (messbarer Raum)

Ist \mathcal{A} eine σ -Algebra über Ω , so nennen wir (Ω, \mathcal{A}) einen *messbaren Raum*.

Definition 4.2.4 (Endlich-additive Mengenfunktion)

Eine *endlich-additive Mengenfunktion* auf (Ω, \mathcal{A}) ist eine Funktion $P : \mathcal{A} \rightarrow [0, 1]$ mit folgenden Eigenschaften:

- i) $P(\Omega) = 1$.
- ii) Für zwei disjunkte Mengen $A, B \in \mathcal{A}$ gilt:

$$P(A \cup B) = P(A) + P(B).$$

Bemerkung 4.2.5

Sind $A_1, \dots, A_n \in \mathcal{A}$ endlich viele paarweise disjunkte Ereignisse, so folgt per Induktion

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

Definition 4.2.6 (Wahrscheinlichkeitsmaß, σ -Additivität)

Ein *Wahrscheinlichkeitsmaß* auf (Ω, \mathcal{A}) ist eine Funktion $P : \mathcal{A} \rightarrow [0, 1]$ mit folgenden Eigenschaften:

- i) $P(\Omega) = 1$.
- ii) Für jede Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ von paarweise disjunkten Mengen gilt

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n).$$

Die Eigenschaft ii) wird *σ -Additivität* eines Wahrscheinlichkeitsmaßes genannt.

Definition 4.2.7 (Wahrscheinlichkeitsraum)

Ist P ein Wahrscheinlichkeitsmaß auf (Ω, \mathcal{A}) , so heißt (Ω, \mathcal{A}, P) *Wahrscheinlichkeitsraum*.

4.2.2 Das Banach-Maß und das skalierte Maß

Die folgenden Definitionen, Sätze, Beweise und Erklärungen stammen, wenn nicht anders angegeben, aus [Rai69, S. 344 ff] und [Hum00, S. 147 ff].

Definition 4.2.8 (Banach-Maß)

Eine auf den Teilmengen von \mathbb{R} definierte reellwertige Mengenfunktion Θ heißt *Banach-Maß*, wenn sie folgende vier Bedingungen erfüllt:

- (a) *Normierung*: $\Theta(\mathbb{R}) = 1$,
- (b) *Positivität*: $\Theta(A) \geq 0$, $A \subset \mathbb{R}$,
- (c) *Endliche Additivität*: $A \cap B = \emptyset \Rightarrow \Theta(A \cup B) = \Theta(A) + \Theta(B)$, $A, B \subset \mathbb{R}$,
- (d) *Translationsinvarianz*: $\Theta(A + s) = \Theta(A)$, $s \in \mathbb{R}$, $A \subset \mathbb{R}$.

Die Menge der Banach-Maße auf \mathbb{R} wird mit \mathcal{T} bezeichnet.

Korollar 4.2.9

Für alle $A, B \subset \mathbb{R}$ gilt:

$$A \subset B \Rightarrow \Theta(A) \leq \Theta(B).$$

Beweis:

Aus der Positivität (b) und der Additivität (c) erhält man:

$$A \subset B \Rightarrow \Theta(A) \leq \Theta(A) + \underbrace{\Theta(B \setminus A)}_{\geq 0} = \Theta(B).$$

□

Definition 4.2.10 (Banach-messbar)

Sei $A \subset \mathbb{R}$. A heißt *Banach-messbar*, wenn die Menge $\{\Theta(A) : \Theta \in \mathcal{T}\}$ einelementig ist, d.h. wenn alle Banach-Maße der Menge A auf \mathbb{R} denselben Wert liefern.

Lemma 4.2.11

Die Menge

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

ist Banach-messbar. Das Banach-Maß von K_a hat den Wert a .

Beweis:

Beweisidee:

Wenn man allein aus den obigen Forderungen an ein Banach-Maß zeigen kann, dass das Maß von K_a den Wert a haben muss, so hat man damit auch gezeigt, dass K_a überhaupt Banach-messbar ist.

Aus der endlichen Additivität (c) und der Translationsinvarianz (d) (vgl. Definition 4.2.8) folgt für $a + b \leq 1$ ($a, b > 0$)

$$\begin{aligned} \Theta(K_{a+b}) &= \Theta\left(\bigcup_{n=-\infty}^{\infty} [n, n+a+b)\right) \\ &= \Theta\left(\bigcup_{n=-\infty}^{\infty} ([n, n+a) \cup [n, n+b) + a)\right) \\ &= \Theta\left(\bigcup_{n=-\infty}^{\infty} [n, n+a) \cup \left(\bigcup_{n=-\infty}^{\infty} [n, n+b)\right) + a\right) \\ &= \Theta(K_a \cup (K_b + a)) \stackrel{(c)}{=} \Theta(K_a) + \Theta(K_b + a) \stackrel{(d)}{=} \Theta(K_a) + \Theta(K_b) \quad (4.2) \end{aligned}$$

Die Beziehung (4.2) lässt sich sofort auf endliche viele Summanden verallgemeinern.

1.Fall: $a = 1$

Mit $K_1 = \mathbb{R}$ und der Normierung (a) erhält man unmittelbar $\Theta(K_1) = 1$ und damit die Behauptung.

2.Fall: $a \in \mathbb{Q}^+$

Mit dem 1. Fall erhält man

$$1 = \Theta(K_1) = \Theta\left(K_{\frac{1}{k} + \frac{1}{k} + \dots + \frac{1}{k}}\right) \stackrel{(4.2)}{=} k \cdot \Theta\left(K_{\frac{1}{k}}\right), \quad \text{für alle } k \in \mathbb{N} \setminus \{0\},$$

woraus sich sofort

$$\Theta(K_{\frac{1}{k}}) = \frac{1}{k}, \quad \text{für alle } k \in \mathbb{N} \setminus \{0\} \quad (4.3)$$

ergibt. Mit Hilfe von (4.2) und (4.3) können wir das Banach-Maß von K_a für alle rationalen Werte von a bestimmen.

Sei $a := \frac{b}{c}$ ($b \in \mathbb{N}$, $c \in \mathbb{N} \setminus \{0\}$). Damit ergibt sich

$$\Theta(K_a) = \Theta(K_{\frac{b}{c}}) \stackrel{(4.2)}{=} b \cdot \Theta(K_{\frac{1}{c}}) \stackrel{(4.3)}{=} \frac{b}{c} = a, \quad \text{für alle } a \in \mathbb{Q}^+. \quad (4.4)$$

3.Fall: $a \in \mathbb{R}^+ \setminus \mathbb{Q}^+$, $0 < a < 1$

Nach dem *Intervallschachtelungsprinzip* gibt es eine Folge von Intervallen mit rationalen Endpunkten $([a_n, b_n])_{n \in \mathbb{N} \setminus \{0\}}$ mit innerstem Punkt a . Die rationale Folge $(a_n)_{n \in \mathbb{N} \setminus \{0\}}$ mit $a_n \in \mathbb{Q}^+$ ist monoton steigend mit $a_n \nearrow a$ und die rationale Folge $(b_n)_{n \in \mathbb{N} \setminus \{0\}}$ mit $b_n \in \mathbb{Q}^+$ ist monoton fallend mit $b_n \searrow a$.

Mit $a_n \nearrow a$ und $b_n \searrow a$ ($a_n, b_n \in \mathbb{Q}^+$) gilt

$$a_n \leq a \leq b_n$$

und damit

$$K_{a_n} \subset K_a \subset K_{b_n}.$$

Mit Korollar 4.2.9 erhält man schließlich

$$\Theta(K_{a_n}) \leq \Theta(K_a) \leq \Theta(K_{b_n}). \quad (4.5)$$

Aus dem 2. Fall ergibt sich $\Theta(K_{a_n}) = a_n$ und $\Theta(K_{b_n}) = b_n$. Damit erhält man unmittelbar mit (4.5)

$$a_n \leq \Theta(K_a) \leq b_n, \quad \text{für alle } n \in \mathbb{N} \setminus \{0\}.$$

Mit $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = a$ folgt

$$\Theta(K_a) = a$$

für alle irrationalen a .

Damit haben wir $\Theta(K_a) = a$ für alle reellen a mit $0 < a \leq 1$ gezeigt. \square

Satz 4.2.12

Für alle $t \in \mathbb{R}$ ist das Intervall $(-\infty, t)$ *nicht* Banach-messbar.

Beweis:

Verwende Theorem 1 aus [Rai57, S. 1033] angewendet auf die charakteristische Funktion über dem Intervall $(-\infty, t)$ mit $t \in \mathbb{R}$. \square

Da die Gesamtheit aller möglichen physikalischen Konstanten auch von Raimi als \mathbb{R}^+ angesehen wird und nicht als ganz \mathbb{R} und *Skaleninvarianz* vorausgesetzt werden soll, definiert Raimi ein Maß auf \mathbb{R}^+ , das *skalierte Maß*, das die gewünschte Eigenschaft erfüllt.

Die Abbildung $\log_{10} : \mathbb{R}^+ \rightarrow \mathbb{R}$ bildet die multiplikative Gruppe (\mathbb{R}^+, \cdot) isomorph auf die additive Gruppe $(\mathbb{R}, +)$ ab, d.h. \log_{10} ist bijektiv und „strukturerhaltend“: $\log_{10}(a \cdot b) = \log_{10}(a) + \log_{10}(b)$

„Ein skaliertes Maß auf \mathbb{R}^+ ist also das „Pendant“ zum Banach-Maß auf \mathbb{R} [...]. [D]ie Existenz von skalierten Maßen ist an die Existenz von Banach-Maßen gebunden und umgekehrt.“³⁰

Definition 4.2.13 (Skaliertes Maß)

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

- (a') *Normierung:* $\Psi(\mathbb{R}^+) = 1$,
- (b') *Positivität:* $\Psi(A) \geq 0$, $A \subset \mathbb{R}^+$,
- (c') *Endliche Additivität:* $A \cap B = \emptyset \Rightarrow \Psi(A \cup B) = \Psi(A) + \Psi(B)$, $A, B \subset \mathbb{R}^+$,
- (d') *Skaleninvarianz:* $\Psi(A \cdot s) = \Psi(A)$, $s \in \mathbb{R}^+$, $A \subset \mathbb{R}^+$.

Die Menge der skalierten Maße auf \mathbb{R}^+ wird mit \mathcal{S} bezeichnet.

Bemerkung 4.2.14

Zu jedem Banach-Maß $\Theta \in \mathcal{T}$ auf \mathbb{R} kann ganz einfach das zugehörige skalierte

³⁰[Hum00, S. 149]

Maß $\Psi \in \mathcal{S}$ auf \mathbb{R}^+ durch

$$\Psi(A) \stackrel{\text{def.}}{=} \Theta(\log_{10}(A))$$

definiert werden, mit $A \subset \mathbb{R}^+$.

Beweis:

Es gilt zu zeigen, dass (a') bis (d') für Ψ erfüllt sind, wenn Θ ein beliebiges Banach-Maß ist und somit (a) bis (d) erfüllt (vgl. Definition 4.2.8 und Definition 4.2.13).

zu (a'): $\Psi(\mathbb{R}^+) = \Theta(\log_{10}(\mathbb{R}^+)) = \Theta(\mathbb{R}) \stackrel{(a)}{=} 1$

zu (b'): $\Psi(A) = \Theta(\log_{10}(A)) \stackrel{(b)}{\geq} 0, \quad A \subset \mathbb{R}$

zu (c'): $A \cap B = \emptyset \Rightarrow$

$$\begin{aligned} \Psi(A \cup B) &= \Theta(\log_{10}(A \cup B)) \\ &= \Theta(\log_{10}(A) \cup \log_{10}(B)) \\ &\stackrel{(c)}{=} \Theta(\log_{10}(A)) + \Theta(\log_{10}(B)) \\ &= \Psi(A) + \Psi(B), \quad A, B \subset \mathbb{R}^+ \end{aligned}$$

zu (d'):

$$\begin{aligned} \Psi(A \cdot s) &= \Theta(\log_{10}(A \cdot s)) \\ &= \Theta(\log_{10}(A) + \log_{10}(s)) \\ &\stackrel{(d)}{=} \Theta(\log_{10}(A)) \\ &= \Psi(A), \quad s \in \mathbb{R}^+, A \subset \mathbb{R}^+ \end{aligned}$$

D.h. die Skaleninvarianz von Ψ ist aufgrund der Logarithmusfunktion äquivalent zur Translationsinvarianz von Θ . □

Definition 4.2.15 (Skaliert-messbar)

Sei $A \subset \mathbb{R}^+$. A heißt *skaliert-messbar* auf \mathbb{R}^+ , wenn die Menge $\{\Psi(A) : \Psi \in \mathcal{S}\}$ einelementig ist, d.h. wenn alle skalierten Maße der Menge A auf \mathbb{R}^+ denselben Wert liefern.

Korollar 4.2.16

$A \subset \mathbb{R}^+$ ist skaliert-messbar, wenn $\log_{10}(A)$ Banach-messbar ist.

Beweis:

Folgt aus (dem Beweis zu) Bemerkung 4.2.14. □

Bemerkung 4.2.17

Skalierte Maße können als Wahrscheinlichkeitsmaße interpretiert werden, d.h. wenn eine Teilmenge $A \subset \mathbb{R}^+$ ein skaliertes Maß $\Psi(A)$ hat, so kann dieser Wert $\Psi(A)$ als Wahrscheinlichkeit interpretiert werden, dass eine „zufällig“ gewählte Zahl x in A liegt:

$$P(x \in A) = \frac{\Psi(A)}{\Psi(\mathbb{R}^+)} = \Psi(A).$$

Skalierte Maße sind nach Definition zusätzlich skaleninvariant. Sie können daher als skaleninvariante Wahrscheinlichkeitsmaße auf \mathbb{R}^+ angesehen werden.

Satz 4.2.18 (Gesetz der führenden Ziffern)

Sei

$$D_p := \bigcup_{n=-\infty}^{\infty} [10^n, (p+1)10^n) \subset \mathbb{R}^+, \quad p = 1, \dots, 9,$$

die Menge aller reellen Zahlen (in Dezimaldarstellung) mit Anfangsziffer $\leq p$.

Dann gilt

$$P(x \in D_p) = \log_{10}(p+1).$$

Beweis:

Betrachten wir das skalierte Maß Ψ von D_p , das über ein Banach-Maß Θ definiert ist. Nach Bemerkung 4.2.14 gilt

$$\Psi(D_p) = \Theta(\log_{10} D_p).$$

Des weiteren ist

$$\log_{10}(D_p) = \bigcup_{n=-\infty}^{\infty} [n, \underbrace{n + \log(p+1)}_{\in (0,1]}), \quad (4.6)$$

d.h. die Menge $\log_{10}(D_p)$ entspricht der Menge $K_{\log_{10}(p+1)}$ aus Lemma 4.2.11. Bestimmt man $\Psi(D_p)$, so erhält man unabhängig von der speziellen Wahl von Θ :

$$\begin{aligned} P(x \in D_p) &\stackrel{4.2.17}{=} \Psi(D_p) \stackrel{\text{def.}}{=} \Theta(\log_{10}(D_p)) \\ &\stackrel{(4.6)}{=} \Theta(K_{\log_{10}(p+1)}) \stackrel{4.2.11}{=} \log_{10}(p+1) \end{aligned}$$

und damit die Behauptung. □

Bemerkung 4.2.19

Das Intervall $(0, T]$, mit $T > 0$, dem im Modell von Pinkham das Maß/die Wahrscheinlichkeit $\frac{1}{2}$ zukommt (vgl. Abschnitt 4.1, S. 47), ist in diesem Modell *nicht* skaliert-messbar.

Beweis:

Es gilt

$$\Psi((0, T]) = \Theta((-\infty, \log_{10}(T)]).$$

Das Intervall $(-\infty, \log_{10}(T)]$ ist nach Satz 4.2.12 nicht *Banach-messbar*.

Damit ist $(0, T]$ nach Korollar 4.2.16 nicht *skaliert-messbar*. □

Somit beweist Raimi die Gültigkeit des *Gesetzes der führenden Ziffern*, also *Benfords Gesetz*, ohne dabei die Existenz einer Verteilungsfunktion aller möglichen positiven Konstanten zu fordern und vermeidet damit einen daraus resultierenden „Mittelpunkt“ all dieser Konstanten. Ein solcher kann seiner Meinung nach nämlich nicht existieren.

KAPITEL 5

Die Methode von Hill

Theodore P. Hill (*1943) ein US-amerikanischer Mathematiker und (seit 2003) emeritierter Professor des *Georgia Institutes of Technology* trug maßgeblich zur Weiterentwicklung der Erkenntnisse über das *Benfordsche Gesetz* bei.³¹

Hill störte sich an den bestehenden Methoden, die bisher herangezogen wurden, um das *Benfordsche Gesetz* zu beweisen (vgl. [Hil95b, S. 888]). Dies sind grob zusammengefasst diskrete und kontinuierliche Summationsmethoden (wie etwa bei Benford (vgl. Abschnitt 1.4.2) selbst oder bei Flehinger, die mit kumulativer Mittelwertbildung und dem Cesàro-Mittel arbeitete (vgl. [Fle66])) und Beweise mittels Skaleninvarianz-Hypothesen (vgl. z.B. Kapitel 3 und 4).

Hill kritisierte an diesen Methoden, dass sie sich entweder nur auf die natürlichen Zahlen beziehen und kontinuierliche Daten vernachlässigen, nicht eindeutig sind und die verwendeten Maße nicht σ -additiv sind (wie etwa das Banach-Maß (vgl. Kapitel 4)).

Was allerdings alle diese Methoden gemeinsam haben, ist, dass sie nicht nur im Dezimalsystem funktionieren, sondern sich auch in jede beliebige andere Basis überführen lassen.

³¹vgl. https://de.wikipedia.org/wiki/Theodore_Hill (26.02.2016)

Daher wählte Hill folgenden Ansatz für seinen Beweis:

„The main purpose [...] is to provide a new formulation of the first-digit problem set on the natural assumption of base-invariance, and to prove that there is a unique countably-additive nonatomic base-invariant probability measure on the positive reals. This probability satisfies the Benford’s law as well as the corresponding n th digit laws for all n and all integer bases.“³²

Im folgenden Abschnitt werden zunächst ein paar grundlegende Definitionen gegeben, bevor dann die Mantissen- σ -Algebra und ein Wahrscheinlichkeitsmaß auf dieser eingeführt werden.

5.1 Grundlagen

Die folgenden Definitionen und Sätze stammen, wenn nicht anders angegeben, aus [Tap13].

Sei Ω eine beliebige Menge und bezeichne $\mathcal{P}(\Omega)$ die Potenzmenge von Ω . Eine Teilmenge $\mathcal{A} \subset \mathcal{P}(\Omega)$ der Potenzmenge wird auch als *Mengensystem* bezeichnet.

Definition 5.1.1 (σ -Algebra)

Ein Mengensystem $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt eine σ -Algebra, falls gilt:

- i) $\Omega \in \mathcal{A}$.
- ii) Für jede Teilmenge $A \in \mathcal{A}$ gilt: $A^c \in \mathcal{A}$.
- iii) Für jede Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ gilt: $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Definition 5.1.2

Es sei $G \in \mathcal{P}(\Omega)$ ein Mengensystem. Dann heißt

$$\sigma(G) := \bigcap_{\substack{\mathcal{H} \supset G \\ \mathcal{H} \text{ ist } \sigma\text{-Algebra}}} \mathcal{H}$$

die von G erzeugte σ -Algebra.

³²[Hil95b, S. 888 f]

Bemerkung 5.1.3

$\sigma(G)$ ist die kleinste σ -Algebra über Ω , die das Mengensystem G umfasst.

Definition 5.1.4 (messbarer Raum)

Ist \mathcal{A} eine σ -Algebra über Ω , so nennen wir (Ω, \mathcal{A}) einen *messbaren Raum*.

Definition 5.1.5 (Borel- σ -Algebra über \mathbb{R}^+ , Borel-Menge)

Die *Borel- σ -Algebra* über \mathbb{R}^+ ist definiert durch $\mathcal{B}(\mathbb{R}^+) := \sigma(\mathcal{O})$, wobei \mathcal{O} das System der offenen Teilmengen von \mathbb{R}^+ bezeichnet. Eine Menge $B \in \mathcal{B}(\mathbb{R}^+)$ wird als *Borel-Menge* bezeichnet.

Bemerkung 5.1.6

Die *Borel- σ -Algebra* über \mathbb{R}^+ ist die kleinste σ -Algebra über \mathbb{R}^+ , die alle offenen Teilmengen umfasst.

Satz 5.1.7

Es gelten folgende Aussagen:

- a) Jede offene Teilmenge O ist eine Borel-Menge.
- b) Jede abgeschlossene Menge A ist eine Borel-Menge.
- c) Jede kompakte Menge K ist eine Borel-Menge.
- d) Jede höchstens abzählbare Menge ist eine Borel-Menge.
- e) Intervalle der Form (a, b) , $[a, b]$, $(a, b]$, $[a, b)$ mit $a < b$ sind Borel-Mengen.
- f) Intervalle der Form (a, ∞) , $[a, \infty)$ mit $a \in \mathbb{R}^+$ sind Borel-Mengen.

Beweis:

Für einen Beweis siehe [Tap13, S. 48 f]. □

Definition 5.1.8 (Wahrscheinlichkeitsmaß, σ -Additivität)

Ein *Wahrscheinlichkeitsmaß* auf (Ω, \mathcal{A}) ist eine Funktion $P : \mathcal{A} \rightarrow [0, 1]$ mit folgenden Eigenschaften:

- i) $P(\Omega) = 1$.
- ii) Für jede Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ von paarweise disjunkten Mengen gilt

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n).$$

Die Eigenschaft ii) wird σ -Additivität eines Wahrscheinlichkeitsmaßes genannt.

Bemerkung 5.1.9

Der Begriff des Borelmaßes wird in der Literatur nicht einheitlich verwendet. Hier wird unter *Borelmaß* bzw. *Borel-Wahrscheinlichkeitsmaß* ein beliebiges Maß bzw. Wahrscheinlichkeitsmaß P auf der Borel- σ -Algebra \mathcal{B} verstanden, d.h.

$$P : \mathcal{B} \rightarrow [0, 1], A \mapsto P(A).$$

Definition 5.1.10 (Wahrscheinlichkeitsraum)

Ist P ein Wahrscheinlichkeitsmaß auf (Ω, \mathcal{A}) , so heißt (Ω, \mathcal{A}, P) *Wahrscheinlichkeitsraum*.

Definition 5.1.11 (Maß)

Ein *Maß* auf (Ω, \mathcal{A}) ist eine Funktion $\mu : \mathcal{A} \rightarrow [0, \infty]$ mit folgenden Eigenschaften:

- i) $\mu(\emptyset) = 0$.
- ii) Für jede Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ von paarweise disjunkten Mengen gilt

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n).$$

Bemerkung 5.1.12

Jedes Wahrscheinlichkeitsmaß ist auch ein Maß, und umgekehrt ist ein Maß genau dann ein Wahrscheinlichkeitsmaß, wenn $\mu(\Omega) = 1$ ist.

Definition 5.1.13

Ist μ ein Maß auf (Ω, \mathcal{A}) , so heißt $(\Omega, \mathcal{A}, \mu)$ ein *Maßraum*.

Definition 5.1.14 (Lebesgue-Maß auf \mathbb{R})

Das Maß λ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, definiert durch

$$\lambda([a, b]) = b - a,$$

für $a, b \in \mathbb{R}$ und $b > a$, wird *Lebesgue-Maß* auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ genannt.

Definition 5.1.15 (Diracmaß³³)

Sei (Ω, \mathcal{A}) ein messbarer Raum. Zu jedem Punkt $a \in \Omega$ wird die zugehörige Abbildung δ_a definiert, die jeder Menge $A \in \mathcal{A}$ den Wert 1 zuordnet, wenn sie a enthält und den Wert 0, wenn sie a nicht enthält:

$$\delta_a(A) = \begin{cases} 1, & \text{falls } a \in A \\ 0, & \text{sonst} \end{cases}$$

Die Abbildung $\delta_a : \mathcal{A} \rightarrow [0, 1]$ wird *Diracmaß* oder *Punktmaß im Punkt x* genannt.

Definition 5.1.16 (μ -Atom, atomlos (vgl. [Kus14, S. 199 f]))

Ist $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, so heißt $A \in \mathcal{A}$ ein μ -Atom, wenn $\mu(A) > 0$ ist und wenn für jedes $B \in \mathcal{A}$, $B \subset A$ gilt $\mu(B) = 0 \vee \mu(A \setminus B) = 0$. Gibt es keine μ -Atome, so nennt man μ *atomlos*.

Beispiel 5.1.17 (vgl. [Kus14, S. 200])

Das Lebesgue-Maß λ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ist atomlos.

Definition 5.1.18 (Invarianz eines Maßes (vgl. [Hil95b, S. 893]))

Ein Maß μ auf (Ω, \mathcal{A}) ist *invariant* unter einer Abbildung $T : \Omega \rightarrow \Omega$, wenn

$$\mu(E) = \mu(T^{-1}(E)) \quad \text{für alle } E \in \mathcal{A}.$$

Notation 5.1.19 (vgl. [Hil95b, S. 889])

Sei $a \in \mathbb{R}$ und $E \subset \mathbb{R}^+$.

³³vgl. <https://de.wikipedia.org/wiki/Diracmaß> (03.03.2016)

- \mathcal{B} bezeichnet die *Borel- σ -Algebra* auf \mathbb{R}^+ .
- $\mathcal{B}(E)$ bezeichnet die *Menge der Borel-Teilmengen* von E .
- $aE := \{ae \mid e \in E\}$
- $a + E := \{a + e \mid e \in E\}$
- \cup symbolisiert die disjunkte Vereinigung von Mengen.

5.2 Die Mantissen- σ -Algebra

Definition 5.2.1 (Mantisse, Basis (vgl. [Glü14, S. 7]))

Jede positive reelle Zahl $x \in \mathbb{R}^+$ kann dargestellt werden als

$$x = m_b \cdot b^n,$$

wobei $n \in \mathbb{Z}$ ist.

Dabei heißt die natürliche Zahl $b \in \mathbb{N}$, $b > 1$, *Basis* und die positive reelle Zahl $m_b \in [1, b)$ *Mantisse* zur Basis b einer positiven reellen Zahl x .

Beispiel 5.2.2

Für $x := 1234 = 1,234 \cdot 10^3$ ist $m_{10} = 1,234$ und für $y := 0.1234 = 1,234 \cdot 10^{-1}$ ist $m_{10} = 1,234$.

Bemerkung 5.2.3 (vgl. [Glü14, S. 8])

Eine negative reelle Zahl $-x$ kann dargestellt werden als

$$-x = -m_b \cdot b^n,$$

wobei $n \in \mathbb{Z}$ ist.

Im Folgenden werden nur Mantissen m_b von positiven reellen Zahlen betrachtet.

Bemerkung 5.2.4 (vgl. [Glü14, S. 8])

Da die Mantisse m_b auf dem Intervall $[1, b)$ definiert ist, ist die Null als führende Ziffer automatisch ausgeschlossen. Erst ab der zweiten Ziffer kann die Null vorkommen.

Definition 5.2.5 (Mantissenfunktion (vgl. [Glü14, S. 8]))

Sei $b \in \mathbb{N}$, $b > 1$ eine Basis. Die Funktion

$$M_b : \mathbb{R}^+ \rightarrow [1, b), x \mapsto m_b,$$

die jeder Zahl $x \in \mathbb{R}^+$ ihre Mantisse $M_b(x) = m_b \in [1, b)$ zuordnet, heißt *Mantissenfunktion* zur Basis b .

Beispiel 5.2.6

Es gilt:

$M_{10}(9) = 9 = M_{100}(9)$ und $M_2(9) = \frac{9}{8}$, da $9 = \frac{9}{8} \cdot 2^3$ und $\frac{9}{8} \in [1, 2)$ ist.

Definition 5.2.7 (Funktion der signifikanten Ziffer (vgl. [Glü14, S. 8 f]))

Sei $b \in \mathbb{N}$, $b > 1$ eine Basis. Die Funktion

$$D_n^{(b)} : \mathbb{R}^+ \rightarrow \{0, 1, 2, \dots, b-1\}, x \mapsto d_n^{(b)}$$

heißt *Funktion der signifikanten Ziffer* zur Basis b .

Dabei bezeichnet $(d_n^{(b)})_{n \in \mathbb{N} \setminus \{0\}}$ die eindeutig bestimmte Folge, für die $d_1^{(b)} \in \{1, 2, \dots, b-1\}$ und $d_n^{(b)} \in \{0, 1, 2, \dots, b-1\}$ ist, für $n \in \mathbb{N}$, $n > 1$, und für die

$$M_b(x) = \sum_{i=1}^n \frac{d_i^{(b)}}{b^{i-1}}$$

ist.

Definition 5.2.8 (n -te signifikante Ziffer (vgl. [Glü14, S. 9]))

$D_n^{(b)}(x)$ heißt *n -te signifikante Ziffer* von x zur Basis $b \in \mathbb{N}$, $b > 1$, kurz $D_n^{(b)}$.

Beispiel 5.2.9

$D_1^{(10)}(1234) = D_1^{(10)}(0.1234) = 1$, $D_2^{(10)}(1234) = D_2^{(10)}(0.1234) = 2$,

$D_3^{(10)}(1234) = D_3^{(10)}(0.1234) = 3$ und $D_4^{(10)}(1234) = D_4^{(10)}(0.1234) = 4$.

Die Mantissen- σ -Algebra \mathcal{M}_b

Der erste Schritt auf dem Weg zu einem geeigneten Wahrscheinlichkeitsmaß ist die Definition einer passenden σ -Algebra. Ein geeigneter Kandidat dafür ist eine Sub- σ -Algebra der Borel- σ -Algebra auf den positiven reellen Zahlen (vgl. [Hil95b, S. 889]).

Sei nun $b \in \mathbb{N}$, $b > 1$ eine feste Basis.

Die Menge der Zahlen, deren erste signifikante Ziffer 1 ist, kann folgendermaßen dargestellt werden (vgl. [Glü14, S. 10]):

$$\{x \in \mathbb{R}^+ \mid D_1^{(b)}(x) = 1\} = \{D_1^{(b)} = 1\} = \bigcup_{n \in \mathbb{Z}} [1, 2) \cdot b^n$$

Allgemein erhält man die Menge

$$M_b^{-1}(E) = \bigcup_{n \in \mathbb{Z}} E \cdot b^n$$

mit $E \in \mathcal{B}([1, b))$, in der alle positiven reellen Zahlen enthalten sind, deren Mantisse in E liegt.

Notation 5.2.10 (vgl. [Hil95b, S. 889])

Bezeichne $\langle E \rangle_b := M_b^{-1}(E)$, mit $E \in \mathcal{B}([1, b))$.

Beispiel 5.2.11 (vgl. [Hil95b, S. 889])

Es gilt:

- a) $\langle \{1\} \rangle_{10} = \{10^n \mid n \in \mathbb{Z}\}$
- b) $\langle [1, b] \rangle_{10} = \mathbb{R}^+$, für $b \in \mathbb{N}$, $b > 1$

Beweis:

Folgt aus der Definition von M_b (vgl. Definition 5.2.5) bzw. M_b^{-1} . □

Definition 5.2.12 (Mantissen- σ -Algebra (vgl. [Hil95b, S. 889]))

Die von M_b auf \mathbb{R}^+ erzeugte σ -Algebra

$$\begin{aligned} \mathcal{M}_b &= \left\{ \bigcup_{n \in \mathbb{Z}} E \cdot b^n \mid E \in \mathcal{B}([1, b)) \right\} \\ &= \{ \langle E \rangle_b \mid E \in \mathcal{B}([1, b)) \} \end{aligned}$$

heißt die *Mantissen- σ -Algebra* zur Basis b .

Bemerkung 5.2.13 (vgl. [Hil95a, S. 356])

Es gilt

$$A \in \mathcal{M}_b \Leftrightarrow A = \bigcup_{n \in \mathbb{Z}} E \cdot b^n$$

mit $E \in \mathcal{B}([1, b))$.

Beweis:

Folgt aus Definition 5.2.12. □

Im Folgenden soll gezeigt werden, dass es sich bei \mathcal{M}_b tatsächlich um eine σ -Algebra über \mathbb{R}^+ handelt (vgl. Definition 5.2.12):

Beweis:

Sei $b \in \mathbb{N}$, $b > 1$.

- $\mathbb{R}^+ \stackrel{5.2.11 \text{ a)}}{=} \langle [1, b) \rangle_b \in \mathcal{M}_b$
- Sei $A \in \mathcal{M}_b$. Dann ist $A = \bigcup_{n \in \mathbb{Z}} E \cdot b^n$ für ein $E \in \mathcal{B}([1, b))$ nach Bemerkung 5.2.13.

Die Menge $A^c = \mathbb{R}^+ \setminus A = \mathbb{R}^+ \setminus (\bigcup_{n \in \mathbb{Z}} E \cdot b^n) = \mathbb{R}^+ \setminus (M_b^{-1}(E))$ umfasst alle positiven reellen Zahlen, deren Mantisse nicht in E , sondern in $[1, b) \setminus E \in \mathcal{B}([1, b))$ liegt. Daher ist $A^c = M_b^{-1}([1, b) \setminus E) \in \mathcal{M}_b$.

- Seien $A_1, A_2, \dots, A_N \in \mathcal{M}_b$. Dann ist $A_i = \bigcup_{n \in \mathbb{Z}} E_i \cdot b^n$ für ein $E_i \in \mathcal{B}([1, b))$ und für $i \in \{1, 2, \dots, N\}$.

Somit ist

$$\bigcup_{i=1}^N \bigcup_{n \in \mathbb{Z}} E_i \cdot b^n = \bigcup_{n \in \mathbb{Z}} \left(\bigcup_{i=1}^N E_i \right) \cdot b^n \in \mathcal{M}_b,$$

da $\bigcup_{i=1}^N E_i \in \mathcal{B}([1, b))$ ist.

Damit ist \mathcal{M}_b nach Definition 5.1.1 eine σ -Algebra über \mathbb{R}^+ . □

Das nächste Lemma legt nahe, dass es sich bei der Mantissen- σ -Algebra um eine geeignetere σ -Algebra für das Wahrscheinlichkeitsmaß der führenden Ziffern handelt als die größere Borel- σ -Algebra.

Lemma 5.2.14 (vgl. [Hil95b, S. 890])

Die σ -Algebra \mathcal{M}_b wird durch $\{D_b^{(i)} \mid i \in \mathbb{N}\}$ erzeugt.

Beweis:

Folgt damit, dass die Mantissenfunktion M_b in Termen der Funktion $D_b^{(i)}$ ausgedrückt werden kann und umgekehrt. \square

Lemma 5.2.15 (vgl. [Hil95b, S. 889])

Für alle $b \in \mathbb{N}$, $b > 1$, gilt:

- i) $\langle E \rangle_b = \bigcup_{k=0}^{n-1} \langle b^k \cdot E \rangle_{b^n}$ für alle $n \in \mathbb{N}$ und $E \in \mathcal{B}([1, b])$;
- ii) $\mathcal{M}_b \subsetneq \mathcal{M}_{b^n} \subsetneq \mathcal{B}(\mathbb{R}^+)$;
Beachte: Diese Inklusion ist strikt.

Beweis:

zu i): Sei $n \in \mathbb{N}$ und sei $E \in \mathcal{B}([1, b])$. Dann gilt:

$$\begin{aligned}
 \bigcup_{k=0}^{n-1} \langle b^k \cdot E \rangle_{b^n} &\stackrel{5.2.12}{=} \bigcup_{k=0}^{n-1} \bigcup_{m \in \mathbb{Z}} (b^n)^m \cdot (b^k \cdot E) \\
 &= \bigcup_{k=0}^{n-1} \bigcup_{m \in \mathbb{Z}} b^{nm+k} \cdot E \\
 &= \bigcup_{m \in \mathbb{Z}} \bigcup_{k=0}^{n-1} b^{nm+k} \cdot E \\
 &= \bigcup_{m \in \mathbb{Z}} (b^{nm} \cdot E \cup b^{nm+1} \cdot E \cup \dots \cup b^{nm+(n-1)} \cdot E) \\
 &= \dots \cup (E \cup b \cdot E \cup \dots \cup b^{n-1} \cdot E) \\
 &\quad \cup (b^n \cdot E \cup b^{n+1} \cdot E \cup \dots \cup b^{2n-1} \cdot E) \\
 &\quad \cup (b^{2n} \cdot E \cup b^{2n+1} \cdot E \cup \dots \cup b^{3n-1} \cdot E) \cup \dots \\
 &= \bigcup_{s \in \mathbb{Z}} b^s \cdot E = \langle E \rangle_b
 \end{aligned}$$

Damit folgt die Behauptung für alle $n \in \mathbb{N}$ und $E \in \mathcal{B}([1, b])$.

zu ii): Nach Definition 5.2.12 sind \mathcal{M}_b und \mathcal{M}_{b^n} definiert als

$$\mathcal{M}_b = \{ \langle E \rangle_b \mid E \in \mathcal{B}([1, b]) \}$$

und $\mathcal{M}_{b^n} = \{ \langle E \rangle_{b^n} \mid E \in \mathcal{B}([1, b^n]) \}$

Offensichtlich gilt, dass $\mathcal{M}_b \subset \mathcal{B}(\mathbb{R}^+)$ und $\mathcal{M}_{b^n} \subset \mathcal{B}(\mathbb{R}^+)$ sind. Für Striktheit der Inklusion vergleiche Bemerkung 5.2.18.

Zeige nun noch, dass $\mathcal{M}_b \subset \mathcal{M}_{b^n}$ ist:

Mit i) gilt:

$$\begin{aligned} \mathcal{M}_b &= \left\{ \bigcup_{k=0}^{n-1} \langle b^k \cdot E \rangle_{b^n} \mid E \in \mathcal{B}([1, b]) \right\} \\ &\stackrel{5.2.12}{=} \left\{ \bigcup_{k=0}^{n-1} \bigcup_{m \in \mathbb{Z}} \langle (b^n)^m \cdot b^k \cdot E \rangle \mid E \in \mathcal{B}([1, b]) \right\} \\ &= \left\{ \bigcup_{m \in \mathbb{Z}} \bigcup_{k=0}^{n-1} \langle (b^n)^m \cdot b^k \cdot E \rangle \mid E \in \mathcal{B}([1, b]) \right\} \\ &= \left\{ \bigcup_{m \in \mathbb{Z}} \langle (b^n)^m \bigcup_{k=0}^{n-1} b^k \cdot E \rangle \mid E \in \mathcal{B}([1, b]) \right\} \\ &\subset \left\{ \bigcup_{m \in \mathbb{Z}} \langle (b^n)^m E' \rangle \mid E' \in \mathcal{B}([1, b^n]) \right\} = \mathcal{M}_{b^n} \end{aligned}$$

Für die Striktheit der Inklusion gilt beispielsweise:

$\langle [1, 20] \rangle_{100} \in \mathcal{M}_{100} = \mathcal{M}_{10^2}$, aber $\langle [1, 20] \rangle_{100} \notin \mathcal{M}_{10}$.

Somit ist $\mathcal{M}_b \subsetneq \mathcal{M}_{b^n} \subsetneq \mathcal{B}(\mathbb{R}^+)$. □

Satz 5.2.16 (vgl. [Hil95a, S. 357] und [Hil95b, S. 889])

Für die Mantissen- σ -Algebra \mathcal{M}_b gilt:

- i) Jede nichtleere Menge in \mathcal{M}_b ist unendlich und besitzt Häufungspunkte in 0 und ∞ .
- ii) Endliche Intervalle der Form (c, d) mit $c, d \in \mathbb{R}^+$ sind *nicht* in \mathcal{M}_b enthalten.

Beweis:

zu i): Nach Bemerkung 5.2.13 gilt: (vgl. [Glü14, S. 12])

$$A \in \mathcal{M}_b \Leftrightarrow A = \bigcup_{n \in \mathbb{Z}} E \cdot b^n \quad \text{mit } E \in \mathcal{B}([1, b])$$

D.h. A ist die unendliche Vereinigung von disjunkten Intervallen und damit unendlich. Die Häufungspunkte in 0 bzw. ∞ erhält man mit $n \rightarrow -\infty$ und $n \rightarrow \infty$

zu ii): Die Behauptung folgt mit i). □

Korollar 5.2.17 (vgl. [Hil95b, S. 889])

Das Problem der Existenz eines „Mittelpunktes“ T aller möglichen positiven Konstanten mit $P(x \in (0, T]) = \frac{1}{2}$, wie Raimi bereits bei Pinkham kritisierte (vgl. Abschnitt 4.1) und selbst in seinem Ansatz berücksichtigte (vgl. Abschnitt 4.2), wird auch hier vermieden, da das endliche Intervall $(0, T]$ nach 5.2.16 nicht in \mathcal{M}_b enthalten und damit nicht messbar ist.

Korollar 5.2.18 (vgl. [Hil95b, S. 890])

Die Mantissen- σ -Algebra \mathcal{M}_b ist eine echte Sub- σ -Algebra der Borel- σ -Algebra \mathcal{B} , d.h. \mathcal{M}_b enthält nicht alle Borel-Mengen.

Beweis:

Nach Satz 5.2.16 sind endliche Intervalle *nicht* in \mathcal{M}_b enthalten, sehr wohl aber in \mathcal{B} . Damit folgt die Behauptung. □

Satz 5.2.19 (vgl. [Hil95a, S. 357] und [Hil95b, S. 889 f])

Für die Mantissen- σ -Algebra \mathcal{M}_b gilt:

- i) \mathcal{M}_b ist *selbst-ähnlich*, d.h. für eine Menge $A \in \mathcal{M}_b$ gilt, dass $b^n \cdot A = A$ ist für alle $n \in \mathbb{N}$.
- ii) \mathcal{M}_b ist bezüglich *Skalarmultiplikation* in \mathbb{R}^+ abgeschlossen, d.h. für $s > 0$ und $A \in \mathcal{M}_b$ gilt, dass $s \cdot A \in \mathcal{M}_b$ ist.
- iii) \mathcal{M}_b ist *nicht* bezüglich *skalarer Addition* abgeschlossen, d.h. für $s > 0$ und $A \in \mathcal{M}_b$ gilt i.d.R. *nicht*, dass $s + A \in \mathcal{M}_b$ ist.
- iv) \mathcal{M}_b ist bezüglich *ganzzahliger Wurzeln* abgeschlossen, d.h. für $n \in \mathbb{N}$ und $A \in \mathcal{M}_b$ gilt, dass $A^{\frac{1}{n}} \in \mathcal{M}_b$ ist; \mathcal{M}_b ist aber *nicht* bezüglich *Potenzen* abgeschlossen.

Beweis:

zu i): (nach [Glü14, S. 13]) Da $A \in \mathcal{M}_b$ ist, gilt nach Bemerkung 5.2.13, dass $A = \bigcup_{m \in \mathbb{Z}} E \cdot b^m$ ist, für ein $E \in \mathcal{B}([1, b])$. Somit ist

$$b^n \cdot A = \bigcup_{m \in \mathbb{Z}} E \cdot b^n \cdot b^m = \bigcup_{m \in \mathbb{Z}} E \cdot b^{n+m} \stackrel{k:=n+m}{=} \bigcup_{k \in \mathbb{Z}} E \cdot b^k \quad \text{für } E \in \mathcal{B}([1, b]).$$

Damit folgt $b^n \cdot A = A$ für alle $n \in \mathbb{N}$.

zu ii): (nach [BH15, S. 19]) Sei $A \in \mathcal{M}_b$ und $s \in \mathbb{R}^+$. Da $A \in \mathcal{M}_b$ ist, gilt nach Bemerkung 5.2.13, dass $A = \bigcup_{n \in \mathbb{Z}} E \cdot b^n$ ist, für ein $E \in \mathcal{B}([1, b])$. Mit i) folgt O.B.d.A. $1 < s < b$.

Dann gilt

$$\begin{aligned} s \cdot A &= \bigcup_{n \in \mathbb{Z}} b^n (s \cdot E) = \bigcup_{n \in \mathbb{Z}} b^n ((s \cdot E \cap [s, b]) \cup (\frac{1}{b} s \cdot E \cap [1, s])) \\ &= \bigcup_{n \in \mathbb{Z}} b^n \cdot C \in \mathcal{M}_b, \end{aligned}$$

da $C := ((s \cdot E \cap [s, b]) \cup (\frac{1}{b} s \cdot E \cap [1, s])) \in \mathcal{B}([1, b])$ ist. Mit Bemerkung 5.2.13 folgt, dass $s \cdot A \in \mathcal{M}_b$ ist.

zu iii): (nach [Hil95b, S. 890]) Beispielsweise gilt: $\{\{1\}\}_{10} \in \mathcal{M}_{10}$, aber

$$\{\{1\}\}_{10} + 5 = \{\dots, 5.01, 5.1, 6, 15, 105, \dots\}$$

liegt *nicht* in \mathcal{M}_{10} .

zu iv): (nach [BH15, S. 19]) Intervalle der Form $[1, b^l]$ mit $0 < l < 1$ erzeugen $\mathcal{B}([1, b])$, d.h. $\mathcal{B}([1, b]) = \sigma(\{[1, b^l] \mid 0 < l < 1\})$. Daher reicht es, die Behauptung für den Spezialfall

$$A = \bigcup_{k \in \mathbb{Z}} b^k [1, b^l] \quad \text{für jedes } 0 < l < 1 \in \mathcal{M}_b$$

zu zeigen. Es gilt

$$\begin{aligned} A^{\frac{1}{n}} &= \bigcup_{k \in \mathbb{Z}} b^{\frac{k}{n}} [1, b^{\frac{l}{n}}] = \bigcup_{k \in \mathbb{Z}} b^k \bigcup_{j=0}^{n-1} [b^{\frac{j}{n}}, b^{\frac{j+l}{n}}] \\ &= \bigcup_{k \in \mathbb{Z}} b^k \cdot C \in \mathcal{M}_b, \end{aligned}$$

da $C := \bigcup_{j=0}^{n-1} [b^{\frac{j}{n}}, b^{\frac{j+1}{n}}] \in \mathcal{B}([1, b])$ ist. Mit Bemerkung 5.2.13 folgt, dass $A^{\frac{1}{n}} \in \mathcal{M}_b$ ist. \square

Anmerkung: (vgl. [Hil95a, S. 357] und [Glü14, S. 12 f])

Die Quadratwurzel einer Menge in \mathcal{M}_b kann aus zwei „Teilen“ bestehen. (Ähnliche Aussagen können auch für höhere Potenzen getroffen werden.) Betrachten wir zum Beispiel die Menge

$$A = \{D_1^{(b)} = 1\} = \bigcup_{n=-\infty}^{\infty} [1, 2) \cdot b^n \quad \text{für } b \in \mathbb{N}, b > 1.$$

Für gerade n gilt, dass $b^{\frac{n}{2}} = b^m$ und für ungerade n gilt $b^{\frac{n}{2}} = b^m \cdot b^{\frac{1}{2}}$ mit $m \in \mathbb{Z}$. Dann gilt für die Quadratwurzel von A

$$\begin{aligned} A^{\frac{1}{2}} &= \bigcup_{n=-\infty}^{\infty} [1, \sqrt{2}) \cdot b^{\frac{n}{2}} \\ &= \bigcup_{m=-\infty}^{\infty} [1, \sqrt{2}) \cdot b^m \cup \bigcup_{m=-\infty}^{\infty} [1, \sqrt{2}) \cdot b^m \cdot b^{\frac{1}{2}} \\ &= \bigcup_{m=-\infty}^{\infty} [1, \sqrt{2}) \cdot b^m \cup \bigcup_{m=-\infty}^{\infty} [\sqrt{b}, \sqrt{2b}) \cdot b^m \in \mathcal{M}_b. \end{aligned}$$

Allerdings ist beispielsweise

$$A^2 = \bigcup_{n=-\infty}^{\infty} [1, 4) \cdot b^{2n} \notin \mathcal{M}_b,$$

da die Exponenten von b nur gerade Zahlen sind und sich A^2 somit nicht wie in Bemerkung 5.2.13 darstellen lässt.

Beispiel 5.2.20 (vgl. [BH15, S. 19 f])

Sei $b = 10$ und sei B die Menge der positiven reellen Zahlen mit führender Ziffer kleiner als 3, d.h.

$$B = \{x > 0 \mid D_1^{(10)}(x) < 3\} = \{D_1^{(10)} < 3\} = \bigcup_{k \in \mathbb{Z}} 10^k [1, 3).$$

Dann ist

$$2 \cdot B = \{x > 0 \mid D_1^{(10)}(x) \in \{2, 3, 4, 5\}\} = \bigcup_{k \in \mathbb{Z}} 10^k [2, 6) \in \mathcal{M}_{10}.$$

Außerdem ist

$$\begin{aligned} B^{\frac{1}{2}} &= \{x > 0 \mid M_{10} \in [1, \sqrt{3}) \cup [\sqrt{10}, \sqrt{30})\} \\ &= \bigcup_{k \in \mathbb{Z}} 10^k ([1, \sqrt{3}) \cup [\sqrt{10}, \sqrt{30})) \in \mathcal{M}_{10}. \end{aligned}$$

Aber

$$B^2 = \bigcup_{k \in \mathbb{Z}} 10^{2k} [1, 9) \notin \mathcal{M}_{10},$$

da beispielsweise $[1, 9) \subset B^2$ ist, aber $[10, 90) \not\subset B^2$ ist.

5.3 Wahrscheinlichkeitsmaß und Baseninvarianz

Im vorherigen Abschnitt wurde ein geeigneter messbarer Raum $(\mathbb{R}^+, \mathcal{M}_b)$ festgelegt, um darauf nun ein Wahrscheinlichkeitsmaß P_B zu definieren, das dem *Benfordschen Gesetz* genügt.

Es soll also „die Wahrscheinlichkeit für alle $x \in \mathbb{R}^+$ berechnet werden, deren Mantisse kleiner als m_b ist – also für die Menge $\{x \in \mathbb{R}^+ \mid M_b(x) < m_b\} = M_b^{-1}([1, m_b]) \in \mathcal{M}_b$ für alle $m_b \in [1, b)$ “ ([Glü14, S. 13]).

Mit $M_b^{-1}([1, m_b]) = \bigcup_{n \in \mathbb{Z}} b^n \cdot [1, m_b)$ und $b \in \mathbb{N}$, $b > 1$ erhält man:

Definition 5.3.1 (vgl. [BH15, S. 32] und [Glü14, S. 13])

Die *Benford Wahrscheinlichkeit* P_B ist das eindeutige Wahrscheinlichkeitsmaß auf $(\mathbb{R}^+, \mathcal{M}_b)$, für das gilt

$$P_B(M_b(x) < m_b) = P_B\left(\bigcup_{n \in \mathbb{Z}} b^n \cdot [1, m_b)\right) := \log_b(m_b) \quad \text{für alle } m_b \in [1, b).$$

Definition 5.3.2 (vgl. [Glü14, S. 14])

$P_B(D_n^{(b)} = d_n^{(b)})$ heißt die Wahrscheinlichkeit, dass die n -te *signifikante Ziffer* gleich $d_n^{(b)}$ ist, kurz $P_B(D_n^{(b)} = d)$.

Bemerkung 5.3.3 (vgl. [Glü14, S. 14])

Die Benford-Wahrscheinlichkeiten $P_B(M_b(x) = m_b)$ einer n -ziffrigen Zahl x las-

sen sich äquivalent in der Form $P_B(D_1^{(b)} = d_1^{(b)}, D_2^{(b)} = d_2^{(b)}, \dots, D_n^{(b)} = d_n^{(b)})$ schreiben.

Definition 5.3.4 (Verallg. Gesetz der führenden Ziffern (vgl. [BH15, S. 9] und [Glü14, S. 14]))

Die Benford-Wahrscheinlichkeit für die *gemeinsame Verteilung der ersten n signifikanten Ziffern* ($n \in \mathbb{N}$) ist gegeben durch

$$\begin{aligned} P_B(D_1^{(b)} = d_1^{(b)}, \dots, D_n^{(b)} = d_n^{(b)}) &= P_B\left(\bigcap_{i=1}^n \{D_i^{(b)} = d_i^{(b)}\}\right) \\ &= \log_b \left(1 + \frac{1}{\sum_{i=1}^n b^{n-i} d_i^{(b)}}\right) \end{aligned}$$

mit $d_1 \in \{1, 2, \dots, b-1\}$ und $d_j \in \{0, 1, \dots, b-1\}$ für $j = 2, \dots, n$.

Nach Hill sollte ein vernünftiges Gesetz über die Verteilung von führenden Ziffern *baseninvariant* sein.

„The underlying motivation for an assumption of base-invariance in a general significant digit law is simply the intuitive feeling that any „reasonable“ such law should be as valid for other integral bases as for base 10.“³⁴

Im Folgenden ist es nun das Ziel, zu zeigen, dass das einzige stetige Wahrscheinlichkeitsmaß auf $(\mathbb{R}^+, \mathcal{M}_b)$, welches *baseninvariant* ist, die *Benford-Wahrscheinlichkeit* P_B ist.

Wenn nicht anders angegeben, stammen die folgenden Definitionen, Sätze und Beweise bzw. Beweisideen aus [Hil95b, S. 890 ff].

Oft ist es hilfreich, ein Wahrscheinlichkeitsmaß auf \mathcal{M}_b mit seiner kanonischen Repräsentation als Wahrscheinlichkeit auf $\mathcal{B}([1, b])$ zu identifizieren (vgl. [Hil95b, S. 890]); daher folgendes Lemma.

³⁴[Hil95b, S. 891]

Lemma 5.3.5

Die Beziehung

$$P(\langle E \rangle_b) = \hat{P}(E), \quad E \in \mathcal{B}([1, b))$$

definiert eine 1–1 Beziehung (Isomorphismus) zwischen einem Wahrscheinlichkeitsmaß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ und einem Borel-Wahrscheinlichkeitsmaß \hat{P} auf $[1, b)$.

Beweis:

Folgt aus Definition 5.2.12. □

Um die folgende Definition eines baseninvarianten Wahrscheinlichkeitsmaßes auf $(\mathbb{R}^+, \mathcal{M}_b)$ zu motivieren gibt Hill folgendes Beispiel:

Betrachten wir die Menge S der positiven reellen Zahlen ($b = 10$), deren erste führende Ziffer kleiner als 5 ist. Mit Lemma 5.2.15 i) kann S dargestellt werden durch ($n = 2$)

$$S = \langle [1, 5) \rangle_{10} = \langle [1, 5) \rangle_{100} \cup \langle [10, 50) \rangle_{100}.$$

Sei nun $a := \log_{10}(5)$. Dann kann S also auf folgende Arten dargestellt werden:

$$S = \begin{cases} \langle [1, b^a) \rangle_b, & \text{falls } b = 10 \\ \langle [1, b^{\frac{a}{2}}) \rangle_b \cup \langle [b^{\frac{1}{2}}, b^{\frac{1+a}{2}}) \rangle_b, & \text{falls } b = 100 \end{cases}$$

Angenommen ein Wahrscheinlichkeitsmaß P auf \mathcal{M}_b ist *baseninvariant*, dann sollte für die Maße der beiden Darstellungen von S gelten:

$$\hat{P}([1, b^a)) = \hat{P}([1, b^{\frac{a}{2}})) + \hat{P}([b^{\frac{1}{2}}, b^{\frac{1+a}{2}})),$$

wobei \hat{P} wie in Lemma 5.3.5 definiert ist. Ähnliches gilt für höhere Potenzen b^n der Basis b .

Diese Überlegungen führen zu folgender wichtigen Definition.

Definition 5.3.6 (Baseninvarianz)

Ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist *baseninvariant*, wenn das zugehörige Wahrscheinlichkeitsmaß \hat{P} auf $\mathcal{B}([1, b))$ der Beziehung

$$\hat{P}([1, b^a)) = \sum_{k=0}^{n-1} \hat{P}\left([b^{\frac{k}{2}}, b^{\frac{k+a}{2}})\right) \tag{5.1}$$

für alle $n \in \mathbb{N}$ und alle $a \in (0, 1)$ genügt.

Bemerkung 5.3.7

Diese Definition der Baseninvarianz gilt nur für Basen b und deren ganzzahlige Potenzen b^n (vgl. Lemma 5.2.15 i)). Aber sie ist dennoch stark genug, um zu zeigen, dass Invarianz in Bezug auf eine Basis b die Invarianz in Bezug auf jede andere Basis impliziert (vgl. Satz 5.3.11).

Beispiel 5.3.8

Das Benford Wahrscheinlichkeitsmaß P_B ist baseninvariant.

Beweis:

Sei P_B wie in Definition 5.3.1 gegeben und sei \hat{P}_B das zu P_B korrespondierende Wahrscheinlichkeitsmaß auf $\mathcal{B}([1, b])$. Sei $a := \log_b(m_b)$. Dann ist

$$\begin{aligned} \sum_{k=0}^{n-1} \hat{P}_B\left(\left[b^{\frac{k}{2}}, b^{\frac{k+a}{2}}\right)\right) &\stackrel{5.3.5}{=} \sum_{k=0}^{n-1} P_B\left(\langle [b^{\frac{k}{2}}, b^{\frac{k+a}{2}}] \rangle_b\right) \\ &= \sum_{k=0}^{n-1} \left(\log_b\left(b^{\frac{k+a}{n}}\right) - \log_b\left(b^{\frac{k}{n}}\right)\right) \\ &= \sum_{k=0}^{n-1} \left(\frac{k+a}{n} - \frac{k}{n}\right) \\ &= \sum_{k=0}^{n-1} \frac{a}{n} \\ &= n \cdot \frac{a}{n} = a = \log_b(m_b) \\ &= P_B(\langle [1, m_b] \rangle_b) = \hat{P}_B([1, b^a]) \end{aligned}$$

Mit Definition 5.3.6 folgt die Behauptung. □

Beispiel 5.3.9

Sei P^* das Wahrscheinlichkeitsmaß auf $(\mathbb{R}^+, \mathcal{M}_b)$, das durch

$$P^*(\langle E \rangle_b) = \begin{cases} 1, & \text{falls } 1 \in E \\ 0, & \text{sonst} \end{cases}$$

für alle $E \in \mathcal{B}([1, b])$ definiert ist.

P^* ist baseninvariant.

Beweis:

Das zu P^* gehörige Borel-Wahrscheinlichkeitsmaß \hat{P}^* auf $[1, b)$ ist das *Diracmaß* δ_1 im Punkt 1 (vgl. Definition 5.1.15):

$$\hat{P}^*(E) = \delta_1(E) = \begin{cases} 1, & \text{falls } 1 \in E \\ 0, & \text{sonst} \end{cases}$$

\hat{P}^* erfüllt offensichtlich die Bedingung für die Baseninvarianz von P^* aus Definition 5.3.6. □

Beispiel 5.3.10

Sei Q_b das Wahrscheinlichkeitsmaß auf $(\mathbb{R}^+, \mathcal{M}_b)$, das durch

$$Q_b(\langle [1, \gamma] \rangle_b) = \frac{\gamma - 1}{b - 1}$$

für $\gamma \in [1, b)$ definiert ist. Q_b ist die *Gleichverteilung* auf \mathcal{M}_b .

Q_b ist *nicht* baseninvariant.

Beweis:

Das zu Q_b gehörige Borel-Wahrscheinlichkeitsmaß \hat{Q}^* auf $[1, b)$ genügt *nicht* der Bedingung (5.1) in Definition 5.3.6:

Sei $n = 2$, $b = 10$, $\gamma = 2$ und sei $a := \log_{10} 2$. Dann gilt

$$\begin{aligned} \sum_{k=0}^1 \hat{Q}^* \left(\left(10^{\frac{k}{2}}, 10^{\frac{k+a}{2}} \right] \right) &= \sum_{k=0}^1 \hat{Q}^* \left(\left(10^{\frac{k}{2}}, 10^{\frac{k+\log_{10} 2}{2}} \right] \right) \\ &= \sum_{k=0}^1 \frac{10^{\frac{k+\log_{10} 2}{2}} - 10^{\frac{k}{2}}}{10 - 1} \\ &= \frac{1}{9} \cdot \underbrace{(\sqrt{2} - 1 + \sqrt{20} - \sqrt{10})}_{>1} \\ &> \frac{1}{9} = \hat{Q}^*([1, 10^a]). \end{aligned}$$

Aus Definition 5.3.6 folgt, dass Q_b *nicht* baseninvariant ist. □

Beim nächsten Satz handelt es sich um den zentralen Satz in Hills Ausführungen, aus dem dann das Gewünschte gefolgert werden kann.

Satz 5.3.11 (Charakterisierung der Baseninvarianz)

Ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist *baseninvariant*, genau dann, wenn es ein $q \in [0, 1]$ gibt, so dass

$$P = qP^* + (1 - q)P_B$$

ist, wobei P_B dem Benford-Wahrscheinlichkeitsmaß entspricht und P^* wie in Beispiel 5.3.9 definiert ist.

Korollar 5.3.12 (vgl. [BH15, S. 76])

Ein stetiges Wahrscheinlichkeitsmaß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist genau dann baseninvariant, wenn $P(E) = P_B(E)$ ist für alle $E \in \mathcal{M}_b$, d.h. wenn P das Benford-Wahrscheinlichkeitsmaß ist.

Beweis:

Folgt unmittelbar aus Satz 5.3.11 mit $q = 0$ und der Tatsache dass P^* unstetig und P_B stetig ist. \square

Bevor Satz 5.3.11 bewiesen werden kann, sind noch ein paar Vorüberlegungen notwendig.

Zunächst wird die Skaleninvarianz eines Wahrscheinlichkeitsmaßes genauer betrachtet.

Definition 5.3.13

Ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist *skaleninvariant* für $s > 0$, falls

$$P(A) = P(sA) \quad \text{für alle } A \in \mathcal{M}_b.$$

Das Wahrscheinlichkeitsmaß P ist auch dann *skaleninvariant*, wenn es skaleninvariant für *ein* $s > 0$ ist, welches *keine* rationale Potenz der Basis b ist, d.h. $s \neq b^n$ für alle $n \in \mathbb{N}$.

Bemerkung 5.3.14

Die Definition der Skaleninvarianz ist sehr allgemein, denn es reicht, die Skaleninvarianz für *einen* skalaren Faktor, der keine rationale Potenz der Basis b ist, vorauszusetzen, um dann mit dem nächsten Satz die Skaleninvarianz für *alle* $s > 0$ zu erhalten (vgl. dazu den Beweis von Satz 5.3.16).

Beispiel 5.3.15

Es gilt:

- a) Das Benford-Wahrscheinlichkeitsmaß P_B ist skaleninvariant.
- b) Das Wahrscheinlichkeitsmaß P^* (vgl. Bsp. 5.3.9) ist *nicht* skaleninvariant.
- c) Das Wahrscheinlichkeitsmaß Q_b (vgl. Bsp. 5.3.10) ist *nicht* skaleninvariant.

Beweis:

zu a): Sei $s > 0$ und sei $m_b \in [1, b)$. Dann gilt

$$\begin{aligned}
 P_B(s \cdot \langle [1, m_b] \rangle_b) &= P_B \left(\bigcup_{k \in \mathbb{Z}} s \cdot b^k \cdot [1, m_b] \right) \\
 &= P_B \left(\bigcup_{k \in \mathbb{Z}} b^k \cdot [s, s \cdot m_b] \right) \\
 &= P_B(s \cdot \langle [s, s \cdot m_b] \rangle_b) \\
 &= \log_b(s \cdot m_b) - \log_b(s) = \log_b \left(\frac{s \cdot m_b}{s} \right) \\
 &= \log_b(m_b) = P_B(\langle [1, m_b] \rangle_b)
 \end{aligned}$$

Mit Definition 5.3.13 folgt die Behauptung.

zu b): Sei $s = 2$. Dann ist $P^*(\langle [1, 3] \rangle_{10}) = 1$, aber $P^*(\langle [2, 6] \rangle_{10}) = 0$.

zu c): Sei $s = 2$. Dann ist $Q_b(\langle [1, 3] \rangle_{10}) = \frac{2}{9}$, aber $Q_b(\langle [2, 6] \rangle_{10}) = \frac{4}{9}$. □

Satz 5.3.16

Das Benford-Wahrscheinlichkeitsmaß P_B ist das eindeutige skaleninvariante Wahrscheinlichkeitsmaß auf $(\mathbb{R}^+, \mathcal{M}_b)$.

Beweis:

P_B ist skaleninvariant nach Beispiel 5.3.15 a). Bleibt noch zu zeigen, dass P_B das *einzigste* skaleninvariante Wahrscheinlichkeitsmaß auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist.

Sei P ein Wahrscheinlichkeitsmaß, das für ein $s > 0$, welches keine rationale Potenz der Basis b ist, skaleninvariant ist. Und sei \bar{P} die b -logarithmische Umskalierung von P auf $\mathcal{B}([0, 1))$, so dass

$$\bar{P}([0, a)) = (\bar{P}([\log_b 1, \log_b b^a])) = \hat{P}([1, b^a)) = P(\langle [1, b^a) \rangle_b) \quad (5.2)$$

für alle $a \in [0, 1)$ ist.

Man sieht, dass eine Skalarmultiplikation bei P mit s einer *irrationalen Rotation* auf dem Kreis $[0, 1) \bmod 1$ bei \bar{P} entspricht. Daher ist die Skaleninvarianz von P bzgl. s äquivalent zu einer Invarianz von \bar{P} bzgl. irrationaler Rotation.

Es ist bekannt, dass das Lebesgue-Maß λ das einzige Maß ist, welches unter irrationaler Rotation invariant ist. Daher muss \bar{P} das Lebesgue-Maß sein. Und mit (5.2) folgt

$$\hat{P}([1, b^a)) = \bar{P}([0, a)) = \lambda([0, a)) = a = \log_b b^a$$

Mit Lemma 5.3.5 folgt, dass $P = P_B$ ist und damit die Behauptung. □

Satz 5.3.17 (Skaleninvarianz impliziert Baseninvarianz)

Wenn ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ skaleninvariant ist, so ist P baseninvariant.

Beweis:

Folgt sofort aus Satz 5.3.11 und Satz 5.3.16. □

Bemerkung 5.3.18

Die Umkehrung von Satz 5.3.17 gilt *nicht*.

Beweis:

Das Wahrscheinlichkeitsmaß P^* ist nach Beispiel 5.3.9 baseninvariant, aber nach Beispiel 5.3.15 b) nicht skaleninvariant. □

Wie im Beweis zu Satz 5.3.16 gesehen, korrespondiert Skaleninvarianz mit einer Invarianz unter irrationaler Rotation T_β , mit $T_\beta : [0, 1) \rightarrow [0, 1)$, $x \mapsto (x + \beta) \bmod 1$, $\beta \in \mathbb{R}^+ \setminus \mathbb{Q}^+$.

Baseninvarianz hingegen korrespondiert mit einer Invarianz unter einer Multiplikation T_n , mit $T_n : [0, 1) \rightarrow [0, 1)$, $x \mapsto nx \bmod 1$, $n \in \mathbb{N}$, $n > 0$. Bei dieser Form der Invarianz sind noch einige grundlegende Fragen ungelöst (vgl. [Hil95b, S. 893]), was auch den Beweis von Satz 5.3.11 schwieriger gestaltet als den Beweis von Satz 5.3.16.

Für den Beweis von Satz 5.3.11 ist noch ein weiteres Lemma notwendig.

Lemma 5.3.19

Ein Borel-Wahrscheinlichkeitsmaß \bar{P} auf $([0, 1), \mathcal{B}([0, 1))$ ist invariant unter der Abbildung $T_n : [0, 1) \rightarrow [0, 1)$, $x \mapsto nx \bmod 1$ für alle $n \in \mathbb{N}$, $n > 0$ genau dann, wenn

$$\bar{P} = q\delta_0 + (1 - q)\lambda$$

ist, für ein $q \in [0, 1]$. δ_0 bezeichnet dabei das Diracmaß im Punkt 0 und λ das Lebesgue-Maß.

Beweis:

Für den Beweis siehe [BH15, S. 77 ff]. □

Korollar 5.3.20

Es folgt sofort aus Lemma 5.3.19:

- Das Lebesgue-Maß ist das einzige *atomlose* Borel-Wahrscheinlichkeitsmaß auf $[0, 1)$, das invariant ist unter der Abbildung T_n , für alle $n \in \mathbb{N}$, $n > 1$.
- Das Lebesgue-Maß ist das einzige *stetige* Borel-Wahrscheinlichkeitsmaß auf $[0, 1)$, das invariant ist unter der Abbildung T_n , für alle $n \in \mathbb{N}$, $n > 1$.

Schließlich kann nun Satz 5.3.11 bewiesen werden.

Beweis von Satz 5.3.11:

„ \Leftarrow “

Sei $b > 1$. Angenommen es gibt ein $q \in [0, 1]$, so dass $P = qP^* + (1 - q)P_B$ ist. Da P^* und P_B baseninvariant sind (vgl. Beispiel 5.3.8 und 5.3.9), folgt leicht, dass auch P baseninvariant ist.

„ \Rightarrow “

Sei nun P baseninvariant auf $(\mathbb{R}^+, \mathcal{M}_b)$ und sei \bar{P} die b -logarithmische Umskalierung von P auf $\mathcal{B}([0, 1])$, wie in (5.2). Nach der Definition der Baseninvarianz 5.3.6 gilt dann

$$\begin{aligned} \bar{P}([0, a]) &= \hat{P}([1, b^a]) = \sum_{k=0}^{n-1} \hat{P}\left(\left[b^{\frac{k}{n}}, b^{\frac{k+a}{n}}\right]\right) \\ &= \sum_{k=0}^{n-1} \bar{P}\left(\left[\frac{k}{n}, \frac{k+a}{n}\right]\right) \end{aligned}$$

für alle $n \in \mathbb{N}$ und alle $a \in (0, 1]$.

Damit lässt sich zeigen, dass \bar{P} invariant ist unter der Abbildung $T_n : [0, 1] \rightarrow [0, 1]$, $x \mapsto nx \pmod{1}$:

$$\begin{aligned} \bar{P}(T_n^{-1}([0, a])) &= \bar{P}\left(\bigcup_{k=0}^{n-1} \left[\frac{k}{n}, \frac{k+a}{n}\right]\right) \\ &= \sum_{k=0}^{n-1} \bar{P}\left(\left[\frac{k}{n}, \frac{k+a}{n}\right]\right) = \bar{P}([0, a]). \end{aligned}$$

Mit Definition 5.1.18 folgt die Invarianz unter der Abbildung T_n . Aus Lemma 5.3.19 erhält man, dass \bar{P} der Gleichung $\bar{P} = q \cdot \delta_0 + (1 - q)\lambda$ für ein $q \in [0, 1]$ genügt.

Nun lässt sich leicht mit (5.2) zeigen, dass für alle $E \in \mathcal{B}([1, b])$ gilt

$$\begin{aligned} P(\langle E \rangle_b) &= \hat{P}(E) = \bar{P}(\log_b(E)) \\ &= q \cdot \underbrace{\delta_0(\log_b(E))}_{\delta_1(E)} + (1 - q) \cdot \lambda(\log_b(E)) \\ &= q \cdot P^*(\langle E \rangle_b) + (1 - q) \cdot P_B(\langle E \rangle_b). \end{aligned}$$

Also ist $P = q \cdot P^* + (1 - q) \cdot P_B$ für ein $q \in [0, 1]$ und damit folgt die Behauptung. \square

Anwendungen des Benfordschen Gesetzes

„[A]fter years of being regarded as a mathematical curiosity, Benford’s law is now being eyed by everyone from tax inspectors to computer designers – all of whom think it could help them solve some tricky problems with astonishing ease.“³⁵

Anwendung findet das Benfordsche Gesetz in den unterschiedlichsten Bereichen. Im Folgenden soll ein kleiner Ausblick darauf gegeben werden, wo sich dieses Gesetz bereits bewährt hat.

6.1 Aufdeckung von Betrug

Der US-Wissenschaftler *Mark Nigrini* entdeckte, dass die Zahlen in Steuererklärungen und Bilanzen dem Benfordschen Gesetz genügen – unter der Voraussetzung, dass sie nicht gefälscht sind. Er untersuchte rund 200.000 Steuererklärungen und stellte fest, dass tatsächlich fast ein Drittel der aufgeführten Beträge mit einer Eins beginnt. Davon ausgehend entwickelte er eine Software, mit deren Hilfe Datensätze auf Übereinstimmungen und Abweichungen von Benfords Gesetz hin

³⁵[Mat99, S. 27]

überprüft werden können. Dabei werden nicht nur die ersten Ziffern, sondern auch die zweiten untersucht (vgl. [Dwo98, S. 228]). Nigrini bezeichnet diese Technik selbst als „digital analysis“ (vgl. [Mat99, S. 29]).

Die Bedenken, dass Betrüger, wenn sie um das Gesetz wissen, gezielt ihre Bilanzen und Steuererklärungen fälschen könnten und damit das Verfahren an Wirksamkeit verlieren könnte, weist Nigrini zurück. Zum einen sagt er, dass es nicht so einfach sei, Daten so zu manipulieren, dass sowohl die ersten als auch die zweiten Ziffern mit dem Gesetz in Einklang stehen (vgl. [Dwo98, S. 228]), zum anderen wissen die Betrüger nicht genau, wie der Test angewendet wird, quartalsweise oder nach Ressorts gruppiert (vgl. [Mat99, S. 30]). „The problem of fraudsters is that they have no idea what the whole picture looks like until all the data are in. [...] Frauds usually involve just a part of a data set [...]“ (ebd.).

Kommt es bei einem Test zu Abweichungen, ist es noch kein Beweis für einen Betrug. Die Daten sollten aber einer genaueren Untersuchung unterzogen werden (vgl. [Dwo98, S. 228]). Unstimmigkeiten können beispielsweise auch durch gewöhnliches Auf- und Abrunden entstehen (vgl. [Mat99, S. 30]).

Mit der Methode von Nigrini konnten beispielsweise die Bilanzfälschungen des Energiekonzerns *Enron* aufgedeckt werden (vgl. [Nig12, S. 207 ff]). Auch gelang es einem Team von Wissenschaftlern der TU Ilmenau nachzuweisen, dass Griechenland seine Wirtschaftsdaten gefälscht hatte (vgl. [Bea11]). Derartige Tests wurden für alle EU-Staaten durchgeführt. Auch Belgien schnitt bei dem Test schlecht ab (vgl. ebd.).

Auch kann die Benford-Analyse herangezogen werden, um mögliche Wahlmanipulationen aufzudecken. So gelang es beispielsweise *Boudewijn F. Roukema*, Hinweise auf mögliche Fälschungen im Rahmen der Präsidentschaftswahlen im Iran 2009 zu ermitteln (vgl. [Rou09]).

6.2 Anwendung in der Informatik

Aber nicht nur im Rahmen der Aufdeckung von Fälschungen ist Benfords-Gesetz im Einsatz. Auch in der Informatik findet es Anwendung. Beispielsweise kann das Benfordsche Gesetz, nach *Donald Knuth*, zur Optimierung von Algorithmen

für floating point Operationen herangezogen werden (vgl. [Hü07, S. 21]). „Wenn man nämlich weiß, welche Ziffern häufiger als andere führende Ziffern auftreten, so kommen bei Operationen wie Addition oder Multiplikation gewisse Überträge (und damit gewisse Registeroperationen) häufiger vor als andere. Durch geschickte Programmierung und Prozessor-Architektur lässt sich dieser Umstand zur Beschleunigung der Algorithmen ausnutzen“ (ebd.).

Für weitere Ausblicke auf Anwendungen im Bereich der Informatik siehe [BH15, S. 228 f].

6.3 Überprüfung von Prognosemethoden

Nach *Theodore Hill* kann Benfords Gesetz auch als „reality check“ für mathematische Modelle, zum Beispiel im Bereich der Prognose der Bevölkerungsentwicklung, herangezogen werden (vgl. [Mat99, S. 30]). Nigrini zeigte, dass die Einwohnerzahlen von Bezirken mit mehr als 3000 Einwohnern in den USA sehr genau dem Benfordschen Gesetz folgen (ebd.). Daher scheint es möglich, Modelle, die die Populationsentwicklung prognostizieren, zu testen, ob sie ebenfalls Benfordverteilte Zahlen liefern. Andernfalls sollten diese Modelle nochmals überdacht werden (ebd.).

Das Benfordsche Gesetz ist nach Nigrini und Hill natürlich kein Allheilmittel, um Betrügern das Handwerk zu legen oder sämtliche Prognosemodelle zu überprüfen und mit Sicherheit wird es auch missbräuchlich verwendet werden. Aber dennoch sind sie sich sicher, dass es noch viele unentdeckte nützliche Anwendungen dieser Gesetzmäßigkeit gibt (vgl. [Mat99, S. 30]). So sagt Hill:

„For me the law is a prime example of a mathematical idea which is a surprise to everyone – even the experts.“³⁶

³⁶[Mat99, S. 30]

Teil II

Für das Mathemuseum

Exponat für das Mathemuseum

Für das Mathemuseum³⁷ der Universität Passau wurde folgendes Exponat angefertigt, anhand dessen von den Besuchern an verschiedenen alltäglichen Datensätzen die Gültigkeit des Benfordschen Gesetzes empirisch überprüft werden kann.

Abbildung 6.1 zeigt das Exponat im Modell und Abbildung 6.2 das originale Exponat. In Abbildung 6.3 findet sich das zugehörige Informationsblatt und in Abbildung 6.4 ein Beispieldatensatz, der zeigt, dass die 1 und die 2 als führende Ziffern in den Flächenangaben der Bundesstaaten der Vereinigten Staaten sehr viel häufiger vorkommen als die anderen Ziffern.

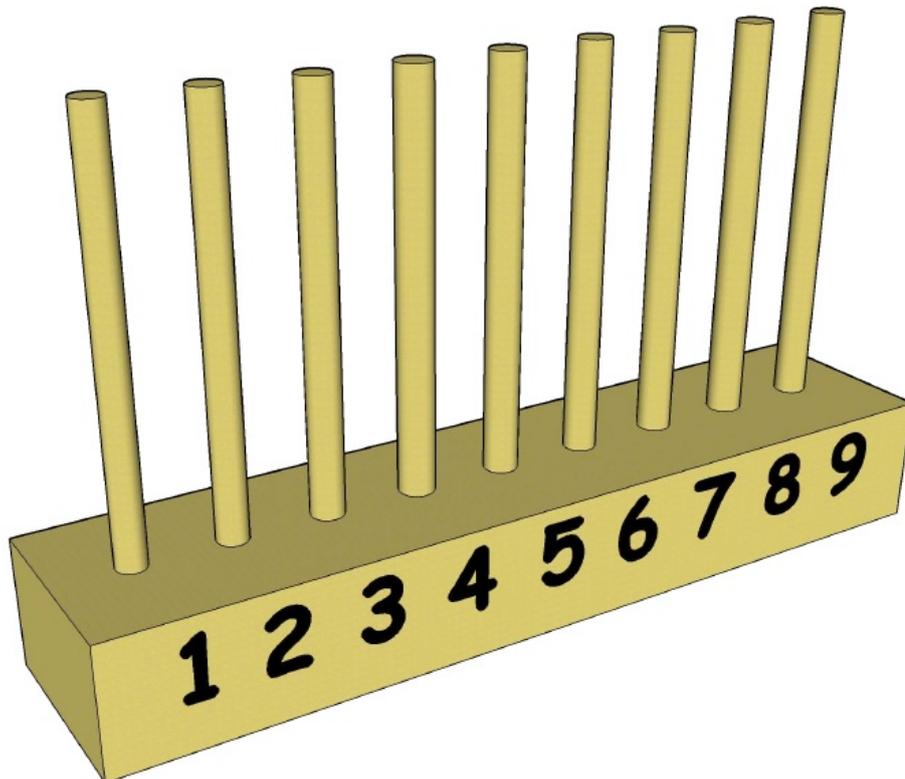


Abbildung 6.1: Modell des Exponats

³⁷Universität Passau, Foyer im Gebäude der Fakultät für Informatik und Mathematik, Innstr. 33; Leitung: Frau Prof. Dr. Brigitte Forster-Heinlein



Abbildung 6.2: Exponat für das Mathemuseum

Das Benfordsche Gesetz

Führende Ziffern kommen *nicht* mit gleicher Wahrscheinlichkeit vor, wie man vielleicht vermuten könnte, sondern unterliegen einer speziellen logarithmischen Verteilung.

Mit **führender Ziffer** ist immer die erste Ziffer einer Zahl gemeint, die ungleich 0 ist.

Zum Beispiel: Die 1 in 1234 bzw. die 1 in 0,01234.

Die Wahrscheinlichkeit, dass die führende Ziffer einer Zahl gleich d ist, lässt sich berechnen durch

$$P(\text{1.Ziffer} = d) = \log_{10} \left(1 + \frac{1}{d} \right),$$

wobei man für d die Ziffern 1, 2, ..., 9 einsetzen kann.

Die nachfolgende Abbildung zeigt die berechneten Wahrscheinlichkeiten für die führenden Ziffern in Prozent. Man sieht, dass **niedrigere Anfangsziffern häufiger vorkommen, als höhere Anfangsziffern.**

Dieses Gesetz der führenden Ziffern wird als **Benfordsches Gesetz** bezeichnet.

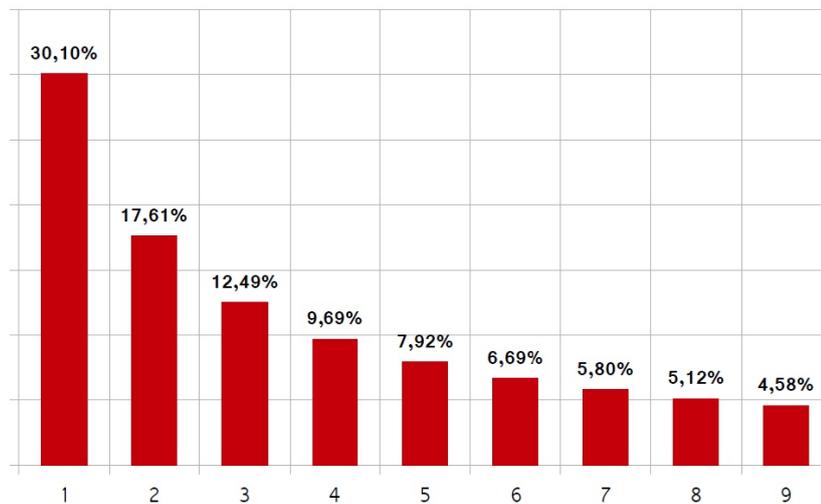


Abbildung 1: Auftretenswahrscheinlichkeiten der führenden Ziffern (in Prozent)

Die Entdeckung des Gesetzes der führenden Ziffern:

Bereits vor Benford machte der Mathematiker **S. Newcomb** im Jahr 1881, diese Entdeckung über die führenden Ziffern. Bei seiner Arbeit mit Logarithmenbüchern stellte er fest, dass die vorderen Seiten stärker abgenutzt waren als die hinteren. Logarithmenbücher sind so aufgebaut, dass vorne die Logarithmen von Zahlen mit niedrigen Anfangsziffern stehen und weiter hinten die der Zahlen mit hohen Anfangsziffern. Er schloss daraus, dass Zahlen mit niedrigen Anfangsziffern häufiger vorkommen müssen als mit hohen Anfangsziffern.

57 Jahre später, im Jahr 1938, machte der Physiker **F. Benford** dieselbe Entdeckung mit den Logarithmenbüchern. Fasziniert von diesem Phänomen sammelte er eine große Menge von Daten aus den unterschiedlichsten Bereichen und bestätigte damit die Existenz einer logarithmischen Verteilungsfunktion für die führenden Ziffern.

Einen exakten mathematischen Beweis lieferte jedoch keiner von Beiden. Der erste Beweis stammt aus dem Jahr 1961 von **R. Pinkham**.

N.	L.	0	1	2	3	4	5	6	7	8	9	P. P.
200	30	103	125	146	168	190	211	233	255	276	298	
201	320	341	363	384	406	428	449	471	492	514	535	22
202	535	557	578	600	621	643	664	685	707	728	749	1
203	750	771	792	814	835	856	878	899	920	941	962	2
204	963	984	1006	1027	1048	1069	1091	1112	1133	1154	1175	3
205	1175	1197	1218	1239	1260	1281	1302	1323	1345	1366	1387	4
206	1387	1408	1429	1450	1471	1492	1513	1534	1555	1576	1597	5
207	1597	1618	1639	1660	1681	1702	1723	1744	1765	1785	1806	6
208	1806	1827	1848	1869	1890	1911	1931	1952	1973	1994	2015	7
209	2015	2035	2056	2077	2098	2118	2139	2160	2181	2201	2222	8
210	2222	2243	2263	2284	2305	2325	2346	2366	2387	2408	2428	9
211	428	449	469	490	510	531	552	572	593	613	634	20
212	634	654	675	695	715	736	756	777	797	818	838	1
213	838	858	879	899	919	940	960	980	1001	1021	1041	2
214	1041	1062	1082	1103	1123	1143	1163	1183	1203	1224	1244	3
215	1244	1264	1284	1304	1324	1344	1364	1384	1404	1424	1444	4

Abbildung 2: Ausschnitt aus einer Logarithmentafel

Arbeitsauftrag:

Anhand der beiliegenden Materialien kann nun empirisch überprüft werden, dass Benfords Gesetz tatsächlich bei sehr vielen alltäglichen Datensätzen Gültigkeit hat. Wählt einen beliebigen Datensatz aus und steckt für jede führende Ziffer (wie oben definiert) einen Ring auf den entsprechenden Stab.

Niedrige Anfangsziffern müssten dabei häufiger vorkommen als hohe Anfangsziffern!

Abbildung 6.3: Informationsblatt für das Mathemuseum

Beispieldatensatz

Bundesstaaten der Vereinigten Staaten



<i>Bundesstaat</i>	<i>Gesamtfläche</i>
Alabama	135.765 km ²
Alaska	1.717.854 km ²
Arizona	295.254 km ²
Arkansas	137.732 km ²
Colorado	269.601 km ²
Connecticut	14.357 km ²
Delaware	6.447 km ²
District of Columbia	177 km ²
Florida	170.304 km ²
Georgia	153.909 km ²
Hawaii	28.311 km ²
Idaho	216.446 km ²
Illinois	149.998 km ²
Indiana	94.321 km ²
Iowa	145.743 km ²
Kalifornien	423.970 km ²
Kansas	213.096 km ²
Kentucky	104.659 km ²
Louisiana	134.264 km ²
Maine	91.646 km ²
Maryland	32.133 km ²
Massachusetts	27.336 km ²
Michigan	250.494 km ²
Minnesota	225.171 km ²
Mississippi	125.434 km ²
Missouri	180.533 km ²

<i>Bundesstaat</i>	<i>Gesamtfläche</i>
Montana	380.838 km ²
Nebraska	200.345 km ²
Nevada	286.351 km ²
New Hampshire	24.216 km ²
New Jersey	22.588 km ²
New Mexico	314.915 km ²
New York	141.299 km ²
North Carolina	139.389 km ²
North Dakota	183.112 km ²
Ohio	116.096 km ²
Oklahoma	181.035 km ²
Oregon	254.805 km ²
Pennsylvania	119.283 km ²
Rhode Island	4.002 km ²
South Carolina	82.932 km ²
South Dakota	199.731 km ²
Tennessee	109.151 km ²
Texas	695.621 km ²
Utah	219.887 km ²
Vermont	24.901 km ²
Virginia	110.785 km ²
Washington	184.665 km ²
West Virginia	62.755 km ²
Wisconsin	169.639 km ²
Wyoming	253.336 km ²

Quelle:
https://de.wikipedia.org/wiki/Liste_der_Bundesstaaten_der_Vereinigten_Staaten_nach_Fläche
(18.06.2016)

Abbildung 6.4: Beispieldatensatz für das Mathemuseum

Literaturverzeichnis

- [Bea11] G. Brähler et al. Wissenschaftler der TU Ilmenau weist Zahlenschummelei Griechenlands nach. <https://www.tu-ilmenau.de/journalisten/presse-meldungen/einzelnachricht/newsbeitrag/8547/>, 2011.
- [Ben38] F. Benford. The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society*, 78(4):551–572, 1938.
- [BGM11] J. Baumeister and T. Garfias Macedo. Von den Zufallszahlen und ihrem Gebrauch. [http://www.math.uni-frankfurt.de/numerik/lehre/Vorlesungen/pros-12/Zufallszahlen und Benford/ran-shell.pdf](http://www.math.uni-frankfurt.de/numerik/lehre/Vorlesungen/pros-12/Zufallszahlen%20und%20Benford/ran-shell.pdf), August 2011.
- [BH15] A. Berger and T.H. Hill. *An Introduction to Benford's Law*. Princeton University Press, New Jersey, 2015.
- [Dia77] P. Diaconis. The distribution of leading digits and uniform distribution mod 1. *The Annals of Probability*, 5(1):72–81, 1977.
- [Dwo98] M. Dworschak. Weiter Weg zur Zwei, Ein kurioses Gesetz der Wahrscheinlichkeitstheorie kann Finanzbeamten helfen, Steuersünder aufzuspüren. *Der Spiegel*, 47:228–229, 1998.

-
- [Fle66] B. J. Flehinger. On the Probability that an Random Integer has Initial D. *The American Mathematical Monthly*, 73(10):1056–1061, 1966.
- [Glü14] M. Glück. *Die Benford-Verteilung – Mathematische Grundlagen und Anwendungsmöglichkeiten in der Marktforschung*. Diplomica Verlag, Hamburg, 2014.
- [Hü07] N. Hüngebühler. Benfords Gesetz über führende Ziffern: Wie die Mathematik Steuersündern das Fürchten lehrt. <http://www.educeth.ch/unt/um/mathe/ana/benford>, März 2007.
- [Hil95a] T. P. Hill. A Statistical Derivation of the Significant-Digit Law. *Statistical Science*, 10(4):354–363, 1995.
- [Hil95b] T. P. Hill. Base-invariance implies Benford’s Law. *Proceedings of the American Mathematical Society*, 123(3):887–895, 1995.
- [Hum00] H. Humenberger. Das Benford-Gesetz - warum ist die Eins als führende Ziffer von Zahlen bevorzugt? *ISTRON-Materialien für einen realitätsbezogenen Mathematikunterricht*, 6:138–150, 2000.
- [Hum08] H. Humenberger. Eine elementarmathematische Begründung des Benford-Gesetzes. *Der Mathematikunterricht*, 54(1):24–34, 2008.
- [HW54] G.H. Hardy and E.M. Wright. *An introduction to the Theory of Numbers*. Oxford Press, Oxford, 1954.
- [KN74] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. John Wiley and Sons, New York, 1974.
- [Kus14] N. Kusolitsch. *Maß- und Wahrscheinlichkeitstheorie*. Springer-Verlag, Wien, 2. edition, 2014.
- [Mat99] R. Matthews. The power of one. *New Scientist*, (10 July 1999) 1999.
- [New81] S. Newcomb. Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics*, 4(1):39–40, 1881.

LITERATURVERZEICHNIS

- [Nig12] M. J. Nigrini. *Benford's law – applications for forensic accounting, auditing, and fraud detection*. Wiley, Hoboken, NJ, 2012.
- [Pin61] R.S. Pinkham. On the distribution of first significant digits. *Annals of Mathematical Statistics*, 32(4):1223–1230, 1961.
- [Rai57] A. R. Raimi. Mean values and Banach limits. *Proceedings of the American Mathematical Society*, 8:1029–1036, 1957.
- [Rai69] R. A. Raimi. On the Distribution of First Significant Figures. *The American Mathematical Monthly*, 76(4):342–348, 1969.
- [Rou09] B. F. Roukema. Benford's law anomalies in the 2009 Iranian presidential election. <http://128.84.21.199/pdf/0906.2789v2.pdf>, 2009.
- [Tap13] S. Tappe. *Einführung in die Wahrscheinlichkeitstheorie*. Springer Verlag, Berlin Heidelberg, 2013.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt worden ist. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet. Die Arbeit wurde nicht bereits in derselben oder einer ähnlichen Fassung an einer anderen Fakultät oder in einem anderen Fachbereich zur Erlangung eines akademischen Grades eingereicht.

.....

Ort, Datum

.....

Unterschrift