

Ziffernanalyse in der Fälschungsaufspürung. Benford´s Gesetz und Steuererklärungen in Theorie und Praxis.

Peter N. Posch*

Abt. Finanzwirtschaft, Universität Ulm
(e-mail: posch@mathematik.uni-ulm.de)

Zusammenfassung

- Theoretische Grundlagen der Ziffernanalyse („Digital Analysis“) werden eingeführt und diskutiert.
- Bekannte statistische Testverfahren zur Fälschungsaufspürung (Fraud Detection) werden diskutiert.
- Eine neue Klasse von Testverfahren, die Invarianztestverfahren, werden eingeführt und praktisch demonstriert.
- Deutsche Steuererklärungen werden mit den vorgestellten Verfahren der Ziffernanalyse untersucht und diskutiert.
- Das Bundesfinanzministerium setzt einige der erörterten Verfahren bereits zur Analyse von Steuererklärungen ein, diese sind jedoch mit Fehlspezifikationen behaftet.

JEL Classification: M49, H20, C49

Key words: Ziffernanalyse, First Digit Law, Fälschungsaufspürung, Steuervermeidung

* Kontakt: Universität Ulm, Abt. Finanzwirtschaft, Helmholtzstrasse 18, 89081 Ulm. Der Autor arbeitet auf dem Gebiet der statistischen Fälschungsaufspürung sowie dem Kreditrisiko, insbesondere dem Rating.

A. Einleitung

Stellen Sie sich vor, man bietet Ihnen folgende Wette an: Sie wählen ein beliebiges Buch mit Zahlentabellen aus, z.B. den Fischer Weltalmanach von 1999. Auf einer zufällig aufgeschlagenen Seite werden die ersten Ziffern jeder dort angegebenen Zahl betrachtet und gezählt, wobei führende Nullen ignoriert werden. Für jede der Ziffer zwischen Eins und Drei erhält Ihr Wettpartner einen Euro, für jede der verbleibenden Ziffern von Vier bis Neun erhalten Sie selbst den gleichen Geldbetrag. Das Angebot wirkt günstig, und man wäre geneigt sich eine Gewinnchance von $\frac{2}{3}$ einzuräumen. – Zufällig wählt man die Seite mit den Spalten 1065/1066. Hier ist die Jahresproduktion ausgewählter Agrarerzeugnisse (Tabak, Tee und Wein) in Mio. Tonnen abgedruckt. Die 54 Einträge der Doppelspalte enthalten nur 19 erste Ziffern zwischen Vier und Neun. Statt des erhofften Gewinns stellt sich ein Verlust von 16 Euro ein.¹ – An Manipulation ist nicht zu denken, waren doch das Buch und die Seite zufällig und selbst gewählt. Man war jedoch intuitiv von einer Gleichverteilung der ersten Ziffern zwischen Eins und Neun ausgegangen. Tatsächlich ergab sich nicht die gleichverteilte relative Häufigkeit von 33.3% für die Ziffern von Eins bis Drei, sondern fast 65%.

Bereits 1881 entdeckte der Astronom Simon Newcomb ein ähnliches Phänomen. Er bemerkte, dass die vorderen Seiten einer Logarithmentabelle weitaus gebrauchter waren als die hinteren.² In einer kurzen Notiz zeigte er, dass „the law of probability of the occurrence of numbers is such that all mantissae of their logarithms are equally probable“ (Newcomb 1881, S. 40). Dies führt – wenn auch nicht direkt ersichtlich – zu einer Häufigkeit von $\log(1 + \frac{1}{d})$ für jede erste Ziffer d zwischen Eins und Neun. Diese Beobachtung blieb weitgehend unbeachtet bis 1938 Frank Benford die gleiche Entdeckung machte. Im Gegensatz zu Newcomb beließ Benford es nicht bei Logarithmentafeln, sondern untersuchte eine Vielzahl verschiedenster Tabellen. Insgesamt zählte er über 20.000 erste Ziffern und fand immer wieder obiges Wettresultat: Die Häufigkeit der ersten signifikanten Ziffern nahm von der Eins mit 30.1%, der Ziffer Zwei mit nur 17.6% bis hin zur Neun mit 4.5% ab. Benford vermutete eine Gesetzmäßigkeit und nannte die Entdeckung „law of anomalous numbers“, wobei Benford bemerkt, dass es sich hierbei um eine Verteilung von Ereignissen handelt, die durch das Medium der Zahl beschrieben werden: „It is not a law of numbers in themselves“ (Benford 1938, S. 554).

Die vorliegende Arbeit beschäftigt sich mit Weiterentwicklungen zu dem mittlerweile „benfordsches Gesetz“ genannten Phänomen der ersten Ziffern, wobei insbesondere die Möglichkeit des Einsatzes zur Fälschungsaufspürung behandelt wird. In *Abschnitt B* wird die Verteilung der Ziffern für beliebige Stellen aus einer Mantissenverteilung kurz hergeleitet, eine gemeinsame Verteilung der ersten k Ziffern angegeben und gezeigt, dass die Abweichung von der Gleichverteilung mit zunehmender Position der Ziffern abnimmt.

In *Abschnitt C* werden Verfahren zur Überprüfung der Anpassungsgüte untersucht und die

üblichen statistischen Anpassungstests auf ihre Einsatzmöglichkeiten in dem sog. Gebiet der Ziffernanalyse überprüft. Ferner werden aus den Eigenschaften der benfordschen Ziffernverteilung neue Testverfahren hergeleitet, die neben ihrer Funktion als Anpassungstest auch eine skeptische, empirische Überprüfung der theoretischen Resultate liefern. Eine der interessantesten Anwendung der Ziffernanalyse stellt die Fälschungsaufspürung dar. Neben theoretischen Überlegungen wird in *Abschnitt D* erstmals die Ziffernanalyse zur Überprüfung deutscher Steuererklärungen benutzt. Für die Durchführung der empirischen Untersuchungen wurde eine Computeranwendung programmiert, welche große Datensätze auf ihr Ziffernverhalten untersucht und neben der Mantissen- und Teststatistikberechnung auch zur Überprüfung der neuen Testverfahren eingesetzt werden kann.

B. Theoretische Grundlagen

Ausgehend von einer stetigen Mantissenverteilung kann die diskrete Verteilung der ersten k Ziffern bei beliebiger Basis b ($b \in \mathbb{N} \setminus \{1\}$) hergeleitet. Hieraus wird die Verteilung an n -ter Stelle und als deren Sonderfall die von Newcomb (1881) und Benford (1938) angegebene Verteilung der ersten Ziffern abgeleitet.³ Probleme treten mit dem Wahrscheinlichkeitsmaß (kurz: W.-Maß) auf, die durch die Definition einer sog. Mantissen- σ -Algebra P jedoch gelöst werden können (siehe Hill (1996)).

Die sig. Ziffern einer positiven reellen Zahl lassen sich mittels der Mantissenfunktion darstellen. Letzter ordnet jeder Zahl $x \in \mathbb{R}^+$ ihre Mantisse $M_b(x) =: m_b \in [1, b[$ zu:

Definition B.1. M_b bezeichnet die *Mantissenfunktion zur Basis b* :

$$\begin{aligned} M_b : \mathbb{R}^+ &\rightarrow [1, b[\\ x &\mapsto m_b \end{aligned}$$

wobei $x = m_b \cdot b^k$ für ein bestimmtes $k \in \mathbb{Z}$ gilt. M ohne Subskript bezeichnet M_{10} .

Beispiel. $M(\pi) = M(0, 31415\dots) = M(31, 415\dots) = \pi$.

Ferner wird die Funktion der signifikanten Ziffer zur Basis b benötigt. Diese ordnet jeder reellen Zahl ihre signifikanten Ziffern zu, wobei führende Nullen ignoriert werden.

Definition B.2 (Hill (1995a), S. 890). Die *Funktion der signifikanten Ziffern zur Basis b* ist gegeben durch:

$$\begin{aligned} D_k^{(b)} : \mathbb{R}^+ &\rightarrow \{0, 1, 2, \dots, b-1\} \\ x &\mapsto d_k^{(b)} \end{aligned}$$

wobei $(d_k^{(b)})_{k \in \mathbb{N}}$ die eindeutig bestimmte Folge bezeichnet, für die $d_1^{(b)} \in \{1, 2, \dots, b-1\}$, $d_k^{(b)} \in \{0, 1, 2, \dots, b-1\}$ für $k \geq 2$ und $M_b(x) = \sum_{i=1}^{\infty} \frac{d_i^{(b)}}{b^{i-1}}$ gilt. Das Superskript bei d_i

wird nur dort angegeben, wo der Kontext keine zweifelsfreie Zuordnung der Basis ermöglicht, hingegen bezeichnet $D_i(\cdot) := D_i^{(10)}(\cdot)$.

Beispiel. $D_4(\pi) = D_4(0, 31415\dots) = D_4(314, 15\dots) = 1$, $D_1^{(2)}(\cdot) \equiv 1$. Die Menge der reellen Zahlen mit erster sig. Ziffer d lautet: $\{x \in \mathbb{R} \mid D_1^{(b)}(x) = d\} = \{D_1^{(b)} = d\} = \bigcup_{n=-\infty}^{\infty} [d, d+1[\cdot b^n$ (vgl. Hill 1996, S. 356).

Definition B.3. Sofern nicht anders angegeben, bezeichnet fortan X eine beliebige Zufallsvariable (kurz: ZV), M_b diejenige ZV, welche die Mantisse zur Basis b repräsentiert, d.i. $M_b := M_b(X)$, und $D_k^{(b)}$ die ZV, welche die k -te signifikante Ziffer von X zur Basis b angibt, d.i. $D_k^{(b)} := D_k^{(b)}(X)$.

Da die Positionen der sig. Ziffern einer Zahl gleich den Positionen der Ziffern ihrer Mantisse sind, kann für die Verteilung der Ziffern eine stetige VF der Mantissen angegeben werden, die eine Verallgemeinerung der Newcomb/Benford-Vermutung darstellt, und aus welcher die Verteilungen der ersten Ziffern hergeleitet wird (Hill 1996, S. 354):

Definition B.4 (Logarithmische Mantissenverteilung). Die *logarithmische VF der Mantissen zur Basis b* wird für alle $m_b \in [1, b[$ gegeben durch $P(M_b \leq m_b) = \log_b(m_b)$.

Mittels Definition B.4 kann nunmehr die gemeinsame Verteilung der ersten k sig. Ziffern ($k \in \mathbb{N}$) angegeben werden:

Theorem B.5 (Hill (1995a), S. 890). Die (logarithmische) gemeinsame Verteilung der ersten k Ziffern ($k \in \mathbb{N}$) ist gegeben durch:

$$P\left(D_1^{(b)} = d_1, D_2^{(b)} = d_2, \dots, D_k^{(b)} = d_k\right) = \log_b \left[1 + \left(\sum_{i=1}^k d_i \cdot b^{k-i} \right)^{-1} \right],$$

wobei $d_1 \in \{1, 2, \dots, b-1\}$ und $d_j \in \{0, 1, 2, \dots, b-1\}$ ($j \geq 2$) gilt.

Beispiel. Für die Zahl 31 ergibt sich demnach: $P(D_1 = 3, D_2 = 1) = \log[1 + 31^{-1}] \approx 0.013788$. Stellt man diese Zahl im Oktalsystem dar, so folgt: $P(D_1^{(8)} = 3, D_2^{(8)} = 7) = \log_8[1 + 31^{-1}]$. Wie vermutet gilt im Dualsystem $P(D_1^{(2)} = 1) = \frac{\ln(2)}{\ln(2)} = 1$.

Sei nun $d \leq M_b(x) < d+1$ mit $d \in \{1, 2, \dots, b-1\}$. Die erste sig. Ziffer in der Darstellung von x zur Basis b ist dann d . Aus der Mantissenverteilung B.4 folgt die von Newcomb (1881) und Benford (1938) angegebene Verteilung der ersten Ziffern sofort als Spezialfall:⁴

Corollary B.6. Die diskrete Dichtefunktion einer ersten Ziffer $d_1 \in \{1, 2, \dots, b-1\}$ ist gegeben durch: $P(D_1^{(b)} = d_1) = \log_b(1 + d_1^{-1})$. Die zugehörige VF lautet: $P(D_1^{(b)} \leq d_1) = \log_b(d_1 + 1)$ (da $P(D_1^{(b)} \leq b-1) = 1$ ist die Dichte wohldefiniert auf $\{1, \dots, b-1\}$).

Die Wahrscheinlichkeit einer Ziffer d_n an n -ter Stelle ergibt sich mittels Satz B.5 als

$$\sum_{\substack{1 \leq d_1 \leq b-1 \\ 0 \leq d_2 \leq b-1 \\ \vdots \\ 0 \leq d_{n-1} \leq b-1}} P\left(D_1^{(b)} = d_1, D_2^{(b)} = d_2, \dots, D_n^{(b)} = d_n\right) .$$

Die Summation kann, nach Feldstein und Goodman (1976), zu folgendem Satz zusammengefasst werden:

Theorem B.7 (Logarithmische Verteilung der n -ten Ziffer). Sei $n > 1, n \in \mathbb{N}$. Die Verteilung der n -ten sig. Ziffer wird für $d_n \in \{0, 1, \dots, b-1\}$ gegeben durch:

$$P(D_n^{(b)} = d_n) = \sum_{k=b^{n-2}}^{b^{(n-1)}-1} \log_b \left(1 + (k \cdot b + d_n)^{-1}\right) .$$

Hieraus folgt sofort die Verteilung der zweiten Ziffern im Dezimalsystem als $P(D_2 = d) = \sum_{k=1}^9 \log(1 + (10 \cdot k + d)^{-1})$. Aufgrund der Sätze B.7 und B.5 folgt, dass die Ziffern der Benford-Verteilung (die Begriffe „Benford-Verteilung“ und „logarithmische (Ziffern)Verteilung“ werden fortan synonym benutzt) an verschiedenen Positionen *nicht* unabhängig sind, so gilt beispielsweise $P(D_1 = 3, D_2 = 1) = \log(1 + 31^{-1}) \approx 0.01379$, jedoch $P(D_1 = 3) \cdot P(D_2 = 1) = \log(1 + 3^{-1}) \cdot (\sum_{k=1}^9 \log(1 + (10 \cdot k + 1)^{-1})) \approx 0.12494 \cdot 0.11389 \approx 0.01423$. Nach Hill (1995b), S. 324 sinkt die Abhängigkeit jedoch mit steigendem Abstand zwischen den betrachteten Ziffern.

Die Abweichung von der a priori vermuteten Gleichverteilung der Ziffern sinkt mit steigendem Abstand von der ersten Ziffer. Beispielsweise liefert das Korollar B.6 für eine führende Eins im Dezimalsystem $\approx 30.103\%$, während mittels Satz B.7 an vierter Stelle nur noch von $\approx 10.014\%$ ausgegangen werden kann.⁵ Die benfordsche Ziffernverteilung an n -ter Stelle konvergiert gegen die Gleichverteilung, die Abweichung von der intuitiven Vorvermutung wird also immer geringer:⁶

Theorem B.8. Für jede feste Basis $b > 1$ gilt: $\lim_{n \rightarrow \infty} P(D_n^{(b)} = d) = \frac{1}{b}, d = 0, 1, \dots, b-1$.

Bemerkung (Feldstein und Goodman (1976)). Trotz des Grenzwertverhaltens nach Satz B.8 und obwohl für festes $n > 1$ der Mittelwert aller n -ten sig. Ziffern $\frac{1}{b}$ ist, gilt $P(D_n^{(b)} = d) = \frac{1}{b}$ für kein n und keine Ziffer $d \in \{0, 1, \dots, b-1\}$.

Anhand dieser Resultate beschränkt sich die praktische Relevanz von Benfords Gesetz auf die ersten vier Stellen einer Zahl im Dezimalsystem. Hiernach ist approximativ von einer Gleichverteilung der Ziffern auszugehen. Mit der Kenntnis der Ziffernverteilung und deren Struktur kann nunmehr folgende Definition einer Benford-ZV gegeben werden. Da insbesondere bei der folgenden empirischen Betrachtung die Dezimalbasis vorhanden sein dürfte, gilt, sofern

nicht anders angegeben $b = 10$. Sämtliche Resultate können jedoch leicht auf $b \in \{2, 3, \dots\}$ erweitert werden.

Definition B.9 (Benford-ZV). Eine ZV X erfüllt Benfords Gesetz, wenn $M_b(X)$ der logarithmischen Mantissenverteilung in Definition B.4 folgt. Eine solche ZV wird als *Benford-ZV* bezeichnet. Eine ZV X erfüllt Benfords Gesetz für die n -te Ziffer, wenn $D_n^{(b)}(X)$ der logarithmischen n -ten Ziffernverteilung in Satz B.7 folgt.

I. Invarianz

Im Folgenden werden zwei Invarianzeigenschaften des Benfordschen Gesetzes kurz erläutert. Der hierzu benötigte Wahrscheinlichkeitsraum und die damit verbundene Definition der Mantissen- σ -Algebra werden hierbei nur angeschnitten. Der interessierte Leser sei auf Hill (1995b) und Hill (1996) verwiesen. Die Definition eines geeigneten W.-Raums hängt eng mit der Struktur von Benfords Gesetz zusammen und eine σ -Algebra sollte so definiert werden, dass das in vorherigem Kapitel betrachtete W.-Maß P σ -additiv ist, und den jeweiligen sig. Ziffern ihre Benford-Wahrscheinlichkeit zuordnet. Da hierbei Aussagen über Ziffern gemacht werden, liegt es nahe als Grundraum \mathbb{R}^+ zu wählen. Man könnte nunmehr vermuten, dass die Borel σ -Algebra auf diesem Grundraum \mathcal{R}^+ geeignet wäre. Jedoch enthält kein einzelnes Intervall alle reellen Zahlen, deren Mantissen zu einer bestimmten Borelmenge gehören. So gilt zum Beispiel im Dezimalsystem, dass jede Zahl aus $[10, 20[$ ihre Mantisse in $[1, 2[$ hat, aber auch jede Zahl aus $[100, 200[$, $[1000, 2000[$ usw. Folglich kann einem einzelnen Intervall nicht alleine durch Betrachtung der korrespondierenden Mantissenmenge eine Wahrscheinlichkeit zugeordnet werden.

Für die Definition der σ -Algebra betrachtet man daher die Menge der positiven reellen Zahlen mit erster Ziffer Eins: $\{D_1^{(b)} = 1\} = \bigcup_{n=-\infty}^{\infty} [1, 2[\cdot b^n$. Verallgemeinert man diese Form, so enthält die Menge

$$S = \bigcup_{n=-\infty}^{\infty} B \cdot b^n \text{ mit Borelmenge } B \subseteq [1, b[\quad (1)$$

alle reellen Zahlen deren Mantisse in B liegt. Da dieser Menge alleine durch Betrachtung der Mantisse eine Wahrscheinlichkeit zugeordnet werden kann, liegt es nahe als σ -Algebra die von den Ziffernfunktionen $D_i^{(b)}$ generierte σ -Algebra zu verwenden (Hill 1995b, S. 323):⁷

Definition B.10 (Hill (1995a), S. 889). Die von den sig. Ziffernfunktionen $D_i^{(b)}$ auf \mathbb{R}^+ generierte σ -Algebra heißt *Mantissen- σ -Algebra zur Basis b* und wird mit \mathcal{M}_b bezeichnet (\mathcal{M} bezieht sich auf die Dezimalbasis). Sie wird gegeben durch:

$$\mathcal{M}_b = \left\{ \bigcup_{n=-\infty}^{\infty} B \cdot b^n, \text{ für alle Borelmengen } B \subseteq [1, b[\right\} .$$

Die Hypothese der Skaleninvarianz entspringt nun der intuitiv plausiblen Idee, dass ein universelles Gesetz der Ziffernverteilung unabhängig von der betrachteten Einheit sein sollte. Wenn beispielsweise Bilanzen in D-Mark ein Ziffernverhalten nach Benford ausweisen, so sollte sich dies bei einer Konvertierung in Euro nicht ändern. Jede Skalierung eines Datensatzes um eine positive reelle Zahl sollte also das Ziffernverhalten des Datensatzes unverändert lassen. Bereits Pinkham (1961) beweist mittels eines klassischen Ansatzes, dass unter allen möglichen Verteilungen erster Ziffern einzig die benfordsche Verteilung diese Eigenschaft erfüllt.⁸

Mittels der Definition der Mantissen- σ -Algebra \mathcal{M}_b nach Hill (1996) kann leicht gezeigt werden, dass die Hypothese der Skaleninvarianz die Ziffernverteilung nach Benford impliziert. Die Skaleninvarianz ist hierbei formal wie folgt definiert:

Definition B.11 (Hill (1995a), S. 892). Ein W.-Maß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ heißt *skaleninvariant* falls für alle $s > 0$ und für alle $S \in \mathcal{M}_b$ gilt: $P(S) = P(sS)$.

Der folgende Satz stellt die Verbindung zu Benfords Gesetz her:⁹

Theorem B.12 (Hill (1995a), S. 893). Ein W.-Maß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist genau dann skaleninvariant, wenn P die logarithmische Verteilung ist.

Die Invarianz bezüglich der Skalierung ist eine erste Eigenschaft des benfordschen Gesetzes, die sich auch zur Konstruktion empirischer Tests eignet. Fraglich ist jedoch, ob die Benford-Verteilung ein Charakteristikum des Dezimalsystems ist. Eine Ziffernverteilung, die von der betrachteten Zahlenbasis abhängt, könnte kaum den Anspruch eines universellen Gesetzes erfüllen und alle bisherigen Beweisansätze unterstützen die Ergebnisse für beliebige Basen $b > 1$.

Zur Motivation der formalen Definition dieser subtileren Hypothese betrachtet man die Menge $S_{(10)}(a) := \bigcup_{n=-\infty}^{\infty} [1, 10^a[\cdot 10^n$ mit $a \in [0, 1[$ (offensichtlich gilt: $S_{(10)}(a) \in \mathcal{M}$ für alle $a \in [0, 1[$). Exakt die gleiche Menge reeller Zahlen kann in der Basis $100 (= 10^2)$ dargestellt werden als $S_{(100)}^{\frac{1}{2}}(a) = \bigcup_{n=-\infty}^{\infty} [1, 100^{\frac{a}{2}}[\cdot 100^n \cup \bigcup_{n=-\infty}^{\infty} [100^{\frac{1}{2}}, 100^{\frac{a+1}{2}}[\cdot 100^n$. Ein baseninvariantes W.-Maß P sollte nun beiden Darstellungen von $S_{(10)}(a)$ den gleichen Wert zuordnen und analog für höhere Basen der Form b^n . Hieraus folgt als natürliche Definition, dass ein baseninvariantes W.-Maß zumindest invariant bezüglich ganzzahliger Wurzeln der Ausgangsbasis sein sollte.¹⁰

Definition B.13 (Hill (1996), S. 358). Ein W.-Maß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ heißt *baseninvariant* falls für alle $m \in \mathbb{N}$ und alle $S \in \mathcal{M}_b$ gilt: $P(S) = P(S^{\frac{1}{m}})$.

Um den Zusammenhang der Baseninvarianz zu Benfords Gesetz herzustellen, betrachtet man die Menge $S_1 := \bigcup_{n=-\infty}^{\infty} \{1\} \cdot b^n \in \mathcal{M}_b$. Diese Menge enthält alle Zahlen mit erster sig.

Ziffer Eins, wobei alle führenden oder folgenden Ziffern Null sind. Der Konstanten Eins kommt insofern eine besondere Bedeutung zu, da ihre Mantisse in allen Basen gleich ist ($M_b(1) = 1, \forall b > 1$). Da die Menge S_1 keine \mathcal{M}_b -messbaren, nicht-leeren Untermengen besitzt, ist das Dirac-Delta-Maß δ_1 auf dieser Menge wohldefiniert.¹¹ Es kann nun eine formale Charakterisierung einer baseninvarianten Ziffernverteilung gegeben werden, wobei P_b das logarithmische W.-Maß bezeichnet:¹²

Theorem B.14 (Hill (1995a), S. 892). Ein W.-Maß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist genau dann baseninvariant, wenn $P = q\delta_1 + (1 - q)P_b$ für ein $q \in [0, 1]$ gilt.

Aus der Herleitung zu Satz B.14 folgt, dass P_b das eindeutige baseninvariante, atomlose W.-Maß auf $(\mathbb{R}^+, \mathcal{M}_b)$ ist. Aus den Sätzen B.12 und B.14 folgt schließlich, dass die Eigenschaft der Skaleninvarianz die der Baseninvarianz impliziert, nicht jedoch umgekehrt, da δ_1 zwar basen-, nicht aber skaleninvariant ist:

Corollary B.15 (Hill (1995a), S. 893). Wenn ein W.-Maß P auf $(\mathbb{R}^+, \mathcal{M}_b)$ skaleninvariant ist, dann ist P baseninvariant.

C. Testverfahren zur Fälschungsaufspürung

Ein praktischer Nutzen der Benford-Verteilung liegt in der Aufspürung von Fälschungen. Hierbei ist besonders die mögliche Kostenreduktion durch ein automatisiertes Verfahren zur Überprüfung der Daten von Vorteil, und so entstand das junge Forschungsgebiet der Ziffernanalyse durch die Erfolge der Überprüfung von Bilanzen mittels Benfords Gesetz. Durch computergestützte Analyse können in kurzer Zeit hierbei große Datenmengen auf ihre Ziffernverteilung hin untersucht und Abweichungen zur näheren Beobachtung markiert werden.

Bevor ein Datensatz auf die Anpassung zur Benford Verteilung getestet wird, sollte überprüft werden, ob die Daten Größen natürlicher Phänomene beschreiben und nicht aus codierten Ziffernfolgen wie ISBN-, Kreditkarten- oder Bankleitzahlen bestehen und die Zahlenwerte des Datensatzes weder von unten noch von oben beschränkt sind. Ein minimaler Zahlenwert von Null ist jedoch akzeptabel. Ferner sollte die Anzahl der *signifikanten* Ziffern Vier übersteigen (Nigrini 2000, S. 24).

I. Anpassungstests

Ist die Anpassung an die Benford-Verteilung ein Indikator für die „Natürlichkeit“ eines Datensatzes, so stellt sich bei empirischen Untersuchungen die Frage nach der erforderlichen Güte der Übereinstimmung. Übliche statistische Anpassungstests können die Hypothese der Benford-Verteilung überprüfen, fraglich ist jedoch, ob deren Macht bei manchen praktischen Anwendungen zu groß ist.

1. Chi-Quadrat-Test

Der χ^2 -Anpassungstest überprüft den Grad der Übereinstimmung zwischen einer Stichprobe vom Umfang n mit einer unbekanntem empirischen VF F_X und einer voll spezifizierten VF F_0 (Gibbons 1985, S. 69-73). Die Nullhypothese lautet $H_0 : F_X(x) = F_0(x)$, für alle x , während die Alternativhypothese $H_1 : F_X(x) \neq F_0(x)$, für mindestens ein x lautet. Da im Allgemeinen relative Häufigkeiten vorliegen dürften, ergibt sich die relevante Teststatistik als:

$$q = n \cdot \sum_{i=1}^9 \frac{(p_i^b - p_i^e)^2}{p_i^e}, \quad (2)$$

wobei p_i^b die beobachtete relative Häufigkeit der i -ten Ziffer und p_i^e analog die erwartete Benford-Häufigkeit bezeichnet. Eine Ablehnung von H_0 erfolgt, falls der kritische Wert $\chi_{1-\alpha; k-1}^2$ der χ^2 -Verteilung mit $k - 1$ Freiheitsgraden und einem Signifikanzniveau von α , kleiner als der Wert der Teststatistik q ist (siehe Tabelle 3 für relevante kritische Werte). Bei den folgenden Untersuchungen werden zusätzlich zu den kritischen Werten die p -Werte angegeben. Der p -Wert – auch empirisches Signifikanzniveau genannt – gibt dasjenige Signifikanzniveau an, bei dem der beobachtete χ^2 -Testwert q gerade am Randpunkt des Verwerfungsbereichs liegt. H_0 wird folglich dann verworfen, wenn $p < \alpha$ gilt (vgl. Bamberg und Baur 1998, S. 213).

Voraussetzung für die Anwendung des χ^2 -Tests ist die *Unabhängigkeit der einzelnen Beobachtungen*. Es sind daher vor der Anwendung Überlegungen bezüglich der Abhängigkeit der Beobachtungen anzustellen bzw. gezielte Tests auf Unabhängigkeit durchzuführen.

Ein bei der Interpretation der Teststatistiken zu beachtendes Problem ist die Sensitivität des χ^2 -Tests bezüglich des Stichprobenumfangs. Während bei einem zu niedrigen Stichprobenumfang q unter H_0 nur approximativ χ^2 verteilt ist, führt ein hoher Stichprobenumfang auch bei nur sehr geringen relativen Abweichungen zu einer Ablehnung der Nullhypothese. Nigrini (2000), S. 78 spricht vom „excess power problem“, welches bei mehr als zehntausend Beobachtungen auftritt und Ley (1996) sieht eine Schwäche der Neyman-Pearson Theorie (vgl. auch die bei Ley (1996) angegebene Literatur). Da die jeweils benötigten Gütekriterien jedoch mittels einer Anpassung des Parameters α – bzw. des zur Annahme der Nullhypothese erforderlichen p -Wertes – realisiert werden können, ist dieser Kritik nicht zu folgen.

2. Kolmogorov-Smirnov-Test

Der Kolmogorov-Smirnov-Einstichprobentest (kurz: KS) basiert ebenfalls auf einem Abstandsmaß zweier Verteilungen, jedoch wird hier das Supremum der absoluten Abweichungen von der empirischen und der theoretischen VF als Teststatistik benutzt:

$$D_n = \sup_x |F_X(x) - F_0(x)|.$$

Die Voraussetzung für den KS-Test ist die *Stetigkeit von $F_0(x)$* . Ist die VF wie bei der Benford-Ziffernverteilung diskret, so ist die Teststatistik D_n nicht parameterfrei. Wenn die unter der Annahme einer stetigen VF berechneten Werte im diskreten Fall benutzt werden, ist das Testniveau höchstens gleich α , der Test ist dann konservativ und von geringerer Güte (Büning und Trenkler 1978, S. 91). Nigrini (2000), S. 78 sieht das Problem der χ^2 - und KS-Tests in der Benford-Analyse in der Macht: „Both [...] test[s] are too powerful [...] to have practical value in an audit setting.“

3. Z-Tests

Die Signifikanz der Abweichung $|p_i^b - p_i^e|$ für jede Ziffer i kann mittels eines Z-Tests (auch: Gauß-Test) berechnet werden als:

$$Z_i = \frac{|p_i^b - p_i^e| \left[-\frac{1}{2n}\right]}{\sqrt{n^{-1} \cdot p_i^e(1 - p_i^e)}}. \quad (3)$$

Eine Stetigkeitskorrektur durch die Subtraktion des Terms $\frac{1}{2n}$ im Zähler findet nur statt, falls $|p_i^b - p_i^e| > \frac{1}{2n}$ ist. Eine signifikante Abweichung liegt dann vor, wenn der Z-Wert $\phi^{-1}(0.975) = 1.96$ überschreitet.

4. Abstandsmaße

Neben den χ^2 - und KS-Tests kann die mittlere absolute Abweichung als Gütekriterium genutzt werden: $MAD := \frac{1}{n} \cdot \sum_{i=1}^n |p_i^b - p_i^e|$. Sie gibt an, wie weit im Mittel die beobachtete von der theoretischen Häufigkeit abweicht. Ferner wird in einem Teil der Literatur die Summe der absoluten Abweichungen, die Summe der quadrierten Abweichungen sowie deren Mittelwert – der mittlere quadratische Fehler – $MSE := \frac{1}{n} \cdot \sum_{i=1}^n (p_i^b - p_i^e)^2$ benutzt. Letztere Maßzahl hat den Vorteil, dass aufgrund der Quadrierung größere Abweichungen stärker gewichtet werden als kleinere.

Da keine kritischen Werte existieren, stellt sich mit der Benutzung von Maßzahlen das Problem der erforderlichen Güte der Anpassung. Ab welchem numerischen Wert soll bei welchem Stichprobenumfang die Nullhypothese abgelehnt werden? Ein erster, unpräziser Lösungsansatz bezieht sich auf praktische Testerfahrung. So gibt Nigrini (2000), S. 118-125 bezüglich der Verteilung der ersten, zweiten und der ersten beiden Ziffern Richtlinien für die Konformität mittels der MAD an (siehe Tabelle 4).¹³

Eine weitere Möglichkeit der Bestimmung kritischer Werte bietet das folgende Vorgehen nach einer Idee von Jamain (2001): Es wird ein Benford-Datensatz D von der Größe des zu untersuchenden Datensatzes – beispielsweise mittels einer geometrischen Reihe – generiert. Hiernach wird D sukzessive mit Nicht-Benford-Daten kontaminiert, wobei an einer zufälligen

Stelle des Datensatzes die gewünschte Anzahl der Daten durch Zufallszahlen einer Normalverteilung mit gleichem Mittelwert und gleicher Varianz wie D ersetzt werden. Geht man beispielsweise von einer 10%igen „Verunreinigung“ des zu untersuchenden Datensatzes aus, so würden genau diese 10% der Daten durch die Zufallszahlen ersetzt werden. Ein Vergleich der so gewonnenen Maßzahlen mit den entsprechenden Abstandsmaßen des zu untersuchenden Datensatzes liefert einen groben Anhaltspunkt für den Grad der Fälschung. Neben der Normalverteilung sind andere Verteilungen denkbar, und die Wahl des gleichen Mittelwerts und der gleichen Varianz dient der Modellierung eines „cleveren“ Fälschers. Es ist fraglich, ob diese Methode zur Gewinnung akkurater Schranken geeignet ist, da zwar die Abstandsmaße mit steigendem Manipulationsgrad ebenfalls größer werden, jedoch gleichzeitig die Varianz der Abstandsmaße – besonders für starke Manipulationen – sehr groß ist.

Wiederum stellt sich jedoch das von Nigrini (2000) angesprochene Problem der Macht, und es kann bei bestimmten Fälschungsanwendungen trotz Ablehnung durch den χ^2 -Test von einer ausreichenden Anpassung zur Benford-Verteilung ausgegangen werden. Legt man hingegen Abstandsmaße als Gütekriterium zu Grunde, bleibt fraglich, wann von einer akzeptablen Anpassung ausgegangen werden kann. Nigrini (1996) versucht mittels einer Maßzahl, dem Verzerrungsfaktor, die Anpassungsgüte zu quantifizieren, und mittlerweile wird nach Angaben Nigrinis dieses Verfahren von den fünf größten Wirtschaftsprüfungsunternehmen sowie von einigen Steuerbehörden zur Aufspürung manipulierter Bilanzen und Steuererklärungen genutzt.

II. Verzerrungsfaktor-Modell

Insbesondere bei Steuererklärungen sind übermäßig viele kleine Zahlen ein Indiz für manipulative Verkleinerung von Beträgen und damit ein Zeichen für mögliche Steuervermeidung. Um den Grad der Verzerrung quantifizieren zu können, entwickelte Nigrini (1996) besagtes Konzept des Verzerrungsfaktors. Im Wesentlichen basiert sein Modell auf dem Vergleich des beobachteten Mittelwertes (BM) mit dem zu erwartenden Mittelwert (EM) einer geometrischen Reihe. Nigrini nimmt an, dass Manipulationen nur innerhalb der gleichen Größenordnung des wahren Eintrags stattfinden, wobei sich „Größenordnung“ auf den Exponenten $k \in \mathbb{N}$ des Intervalls $[1, 10[\cdot 10^{k-1}$ bezieht. Ferner wird angenommen, dass der Manipulationsgrad über alle Größenordnungen gleich hoch ist. Um Eindeutigkeit zu erreichen, werden alle Einträge mit Werten < 10 herausgestrichen und Werte ≥ 100 in den Bereich $[10, 100[$ skaliert. Dieses Vorgehen entspricht – bis auf die Verwerfung von Einträgen < 10 – der Berechnung der dekadischen Mantisse und deren Multiplikation mit dem Faktor 10. Der Verzerrungsfaktor (DF) wird hiernach als relative Abweichung des beobachteten vom erwarteten Mittelwert berechnet: $DF := \frac{BM-EM}{EM}$. Der Wert von EM hängt zwar von der Anzahl der Glieder der

geometrischen Reihe ab, nach Nigrini (1996), S. 76 ist der Einfluss bei mehr als 500 Beobachtungen vernachlässigbar und es kann ungefähr von einem Wert von 39.04 ausgegangen werden.

Allerdings ist kaum ersichtlich, warum eine Restriktion der Daten auf den Bereich $[10, 100[$ in Kauf genommen werden sollte, zumal die Berechnung mittels der Mittelwerte bereits einen Informationsverlust darstellt. Eine denkbare Erweiterung des Modells wäre daher die Berechnung des Mittelwerts der Mantissen eines Datensatzes und dessen Vergleich mit dem Referenzmittelwert einer geometrischen Folge auf $[1, 10[$. Im Folgenden werden die Ergebnisse für den Bereich der dekadischen Mantisse hergeleitet, also eine modifizierte Version von Nigrinis DF-Modell gegeben.

Definition C.1 (Bronstein u. a. (1999), S. 18). Eine Folge (a_n) heißt geometrische Folge, wenn der Quotient zweier aufeinanderfolgender Folgenglieder konstant ist, d.h. wenn gilt: $\frac{a_{n+1}}{a_n} = q = \text{konstant}$, also $a_n = a_1 \cdot q^{n-1}$.

Um eine Skalierung im Bereich $[\alpha, \beta[$ bei N Folgengliedern zu erreichen, setzt man $a_1 := \alpha$ und q so, dass $a_N \leq \beta$ gilt:¹⁴

$$a_N \leq \beta \Leftrightarrow q \leq 10^{\frac{\log(\beta) - \log(\alpha)}{N}} .$$

Skaliert man die Folge nun auf den Bereich der Mantissen $[1, 10[$, so ergibt sich hier $q = 10^{\frac{1}{N}}$ als konstanter Quotient. Der Endpunkt $\beta = 10$ wird approximativ für hinreichend große N erreicht, da $a_N = a_1 \cdot 10^{\frac{N-1}{N}} = a_1 \cdot 10^{1 - \frac{1}{N}}$ gilt.

Für die endliche Summe aller Folgenglieder, die geometrische Reihe, gilt allgemein:

$$S_n := \sum_{i=1}^n a_i = a_1 \cdot \frac{q^n - 1}{q - 1} .$$

Der Mittelwert lautet demnach:

$$\begin{aligned} EM &= \frac{S_N}{N} \\ &= (\text{aufgrund der Skalierung auf } [1, 10[) \\ &= \frac{9}{(10^{\frac{1}{N}} - 1) \cdot N} . \end{aligned}$$

Der Grenzwert des Mittelwertes lautet nun:

$$\lim_{N \rightarrow \infty} \frac{S_N}{N} = \frac{9}{\ln(10)} \approx 3.908650337 .$$

Für den Verzerrungsfaktor Nigrinis $DF := \frac{BM - EM}{EM}$ ergibt sich ein Wertebereich von $[-0.74411, 1.55885[$. Fraglich ist der Wertebereich hinsichtlich der von Nigrini vorgeschlagenen Interpretation als Prozentsatz der Übereinstimmung mit Benfords Gesetz. (Für eine

Herleitung der Varianz und daran anknüpfend eines Konfidenzintervalls siehe Nigrini (1996), S.88f.)

Die Modellerweiterung hingegen wurde bei den folgenden Untersuchungen berücksichtigt und steht als Maßzahl in der elektronischen Analyse zur Verfügung. Es stellt sich jedoch beim DF sowie beim Mantissen-DF (MDF) das gleiche Problem wie bei den anderen Abstandsmaßen, da auch hier nicht deutlich ist, wie groß der Wert von DF sein darf, um noch von einer akzeptablen Anpassung ausgehen zu können.

III. Zifferntests

Nigrini (2000) führt am Beispiel der amerikanischen Volkszählung von 1990 fünf Tests bezüglich der Ziffernverteilung vor. Hierbei werden mehrere Verfeinerungsschritte durchlaufen, um letztlich eine mögliche Fälschungsaufspürung zu erreichen. Als Anpassungskriterium benutzt Nigrini Balkendiagramme und das Abstandsmaß der MAD. Sukzessive werden hierbei Diagramme der ersten Ziffern (neun Möglichkeiten), der zweiten Ziffern (zehn Möglichkeiten), der ersten beiden Ziffern (99 Möglichkeiten) sowie der ersten drei Ziffern (900 Möglichkeiten) betrachtet und mit der Benford-Verteilung graphisch verglichen. Mittels Z-Tests werden Abweichungen von den erwarteten Ziffernhäufigkeiten auf Signifikanz untersucht und zur Beobachtung markiert. Nigrini (2000), S. 42-51 merkt an, dass die Tests auf die ersten und zweiten Ziffern im Rahmen einer Fälschungsaufspürung zu grob sind, da sie eine zu große Datenmenge zur näheren Überprüfung markieren würden. Der fünfte Zifferntest überprüft schließlich die letzten beiden Ziffern (100 Möglichkeiten) und dient der Aufspürung von Rundungsdifferenzen und Zufallszahlen.

IV. Invarianz-Testverfahren

Das Problem der Nichtanwendbarkeit des χ^2 -Tests bei Abhängigkeit der Daten sowie die Unschärfe der Schranken bei Abstandsmaßen machen ein Testverfahren wünschenswert, welches eine klare Aussage über die Anpassung an die logarithmische Verteilung liefert. Neben den üblichen Anpassungstests können die Transformationsinvarianzeigenschaften der Benford-Verteilung als Testverfahren genutzt werden.¹⁵

Die Idee der Invarianz-Tests – welche bereits der Begriff „*Invarianz*“ nahe legt – ist dabei folgende: Bei der Überprüfung der Invarianzeigenschaften werden auf jeder Transformationsstufe – zum Beispiel für jede Basis bei der Baseninvarianz – verschiedene Abstandsmaße berechnet. Bei einer Betrachtung der einzelnen Abstandsmaße bliebe das Problem der erforderlichen Anpassungsgüte bestehen. Daher wird für jeden Abstandsmaßtyp die Varianz sowie der Mittelwert über *alle* Transformationen berechnet. Sofern die Varianz nicht signifikant

von Null verschieden ist, kann von einem Vorliegen der untersuchten Invarianzeigenschaft ausgegangen werden. Ferner wäre ein Vergleich des jeweiligen Mittelwerts der Invarianzuntersuchung mit der entsprechenden Maßzahl des ursprünglichen Datensatzes möglich. Auch hier könnte man bei nicht signifikanter Abweichung von einer Bestätigung der Nullhypothese der jeweiligen Transformationsinvarianz ausgehen. Da die Überprüfung zwar einen höheren Rechenaufwand erfordert, nicht jedoch eine den statistischen Anpassungstest ähnliche Macht hat, ist der praktische Nutzen dieser Testverfahren noch begrenzt. Ob Invarianztests dieser Art dennoch für die Ziffernanalyse ausreichen, müsste in empirischen Untersuchungen mittels eines Vergleichs der üblichen Anpassungstests bestätigt werden. Auch eine theoretische Herleitung der Konfidenzintervalle der jeweiligen Abstandsmaßzahlen für das Vorliegen der entsprechenden Invarianz würde die praktischen Möglichkeiten der Anwendung erweitern.

Im Folgenden werden die spezifischen Möglichkeiten solcher Invarianz-Tests anhand des Datensatzes *CPI* – der die unbereinigten, quartalsweisen Inflationsraten von 1960 bis 2000 enthält und eine gute Anpassung an die logarithmische Verteilung aufweist – gezeigt. Ein „natürlicher“ Datensatz wird einem generierten Datensatz mit perfekter Benford-Übereinstimmung deshalb vorgezogen, da Anhaltspunkte für empirische Schranken gewonnen werden sollen.¹⁶ Es wird jeweils das Ziffernverhalten der ersten Ziffern mittels der Abstandsmaße MAD und MSE, sowie als Vergleichsmaßstab des χ^2 -Tests untersucht.¹⁷ Eine Erweiterung der Untersuchung auf Ziffern an n -ter Stelle oder die Betrachtung der ersten k -Ziffern ist leicht möglich.

1. Skaleninvarianz

Die Benford-Verteilung ist invariant bezüglich positiver Skalierungen. Es kann also ein Datensatz mit positiven Konstanten multipliziert und hiernach die jeweilige Anpassung gemessen werden. Da Auswahl und Anzahl der gewählten Konstanten das Ergebnis beeinflussen, sollte deren Festlegung vor der Untersuchung erfolgen. Es bietet sich für die Wahl der Konstanten das Intervall von $[1, b[$ und eine feste Schrittweite, z.B. 0.1, an.

Auf dem hier gewählten Intervall $[1, 10[$ ergeben sich bei 0.1er Schritten 91 Transformationsstufen. Der Mittelwert der MAD liegt bei 0.01941, die Varianz bei $1.38525 \cdot 10^{-5}$. Für den MSE ergibt sich über die 91 Multiplikationen ein Mittelwert von 0.00063 und eine Varianz von $9.68053 \cdot 10^{-8}$. Der χ^2 -Wert schließlich liegt im Mittel bei 7.48687 (Varianz: 7.48967). Der unskalierte Datensatz befindet sich erwartungsgemäß unter denen mit der besten Anpassung. Die geringe Varianz und die nahe an den Maßzahlen der unskalierten Daten liegenden Durchschnittsmaßzahlen sprechen insgesamt für das Vorliegen von Skaleninvarianz. Die Abbildung 2 verdeutlicht den Verlauf der Anpassung.

2. Baseninvarianz

Ebenfalls eignet sich die – von der Skaleninvarianz implizierte – Eigenschaft der Baseninvarianz als Testmöglichkeit. Diese ergibt sich analog zur Skaleninvarianz aus der Konvertierung der Daten in verschiedene Basen und der basenweisen Bestimmung der Güte der Anpassung. Einschränkend muss angemerkt werden, dass sich für das Testen der Baseninvarianz nur die Basen von Drei bis Neun eignen, da im Binärsystem die Ziffern ab der zweiten Stelle gleichverteilt sind, während für Basen $b > 10$ komplizierte Rechenprozeduren nötig werden.

Die MAD hat über die Basentransformationen einen Mittelwert von 0.02888 und eine Varianz von 0.00023, das Mittel der MSE liegt bei 0.00150, die Varianz bei Null. Der χ^2 -Test hat eine mittlere Teststatistik in Höhe von 6.631 (Varianz: 14.1439). Auch liefern die Abstandsmaße in jeder einzelnen Basis Anhaltspunkte für eine Konformität der Daten zu der logarithmischen Ziffernverteilung (siehe Tabelle 5). Insgesamt kann durch die geringe Varianz und die enge Anpassung in jeder Basis von Baseninvarianz ausgegangen werden. Bei der Interpretation des χ^2 -Wertes ist indes zu beachten, dass es jeweils $b - 2$ Freiheitsgrade gibt. Die Basis Drei weist – da es hier lediglich zwei Klassen gibt – folglich die kleinste χ^2 -Teststatistik auf.

3. Inversionsinvarianz

Das benfordsche Gesetz ist ferner invariant bezüglich der Inversion. Der invertierte Datensatz weist folgende Mittelwerte der Abstandsmaße auf. Die MAD liegt bei 0.01917 (Varianz: $6.5973 \cdot 10^{-6}$), der MSE bei $5.2 \cdot 10^{-4}$ (Varianz: $2.46535 \cdot 10^8$) und der χ^2 -Test bei 6.2011 (Varianz: 4,51807). Alle Maßzahlen waren damit niedriger als die entsprechenden Ergebnisse des Ausgangsdatsatzes. Eine Berechnung der Varianz ist bei nur zwei Werten sehr bedingt aussagekräftig, die Hypothese der Inversionsinvarianz könnte aber dennoch als recht gut verifiziert angesehen werden. Es ist allerdings fraglich, wie akkurat dieser spezielle Vergleich ist und ob entsprechende Ablehnungsschranken hierfür gefunden werden können. Die Einfachheit der Berechnung dieses Test könnte ihn jedoch als nützliches „K.O.“-Kriterium für die Ziffernanalyse erweisen, obwohl unter den hier vorgestellten Invarianz-Tests dieser Test der statistisch schwächste sein dürfte.

D. Steuervermeidung

Es wird nun die Möglichkeit des Einsatzes von Benfords Gesetz zur Aufspürung von Steuervermeidung, -umgehung und -hinterziehung betrachtet. Hierbei stellt sich die Frage, ob manipulative Angaben in Steuererklärungen überhaupt mittels einer Analyse des Ziffernverhaltens entdeckt werden können und welche Arten von Falschangaben hierbei ggf. unentdeckt

bleiben.

Es wird angenommen, dass unmanipulierte Einträge in Steuererklärungen dem benford-schen Gesetz folgen. Psychologische Untersuchungen zeigen, dass Menschen dazu neigen, starke Wiederholungen sowie Extreme zu vermeiden: So lässt Hill (1988) seine Studenten eine sechsstellige Zahl ausdenken und überprüft das Ziffernverhalten der ersten und zweiten Ziffer. Er führt sowohl einen χ^2 - als auch einen KS-Test durch, welche die Hypothese der Benford-Verteilung signifikant ablehnen. Hill (1999) bittet seine Studenten eine Münze wahlweise 200mal tatsächlich zu werfen oder sich die Ergebnisse auszudenken. Bei wirklichen Zufallswürfen ist eine Wiederholung von sechs „Köpfen“ hintereinander sehr wahrscheinlich, bei den „gefälschten“ Ergebnissen kamen diese jedoch nur selten vor. Daher wird ferner angenommen, dass von Personen spontan ausgedachte Zahlen kein solches Ziffernverhalten zeigen.

Diese Hypothesen implizieren, dass mittels der Ziffernanalyse lediglich *ungeplante* Manipulationen entdeckt werden können (Nigrini 1996, S. 77). Geplante Steuerhinterziehung mittels einer langfristigen Manipulation verschiedener Formulare und Dokumente sind nicht aufspürbar. Aufgrund der Skaleninvarianz ist es ferner nicht möglich Datensätze, die mittels einer konstanten Skalierung verändert wurden, von dem Originaldatensatz zu unterscheiden.¹⁸

Eine Untersuchung amerikanischer Steuererklärungen führen erstmals Busta und Sundheim (1992) durch. Sie benutzen einen von der US-Regierung zur Verfügung gestellten Datensatz mit 50 Tausend Einträgen für das Jahr 1982.¹⁹ Die Hypothese der Benford-Ziffernverteilung konnte nicht bestätigt werden. Nigrini (1996) untersucht in dem gleichen Datensatz die Ziffernverteilung der Felder „Erhaltene Zinsen“ und „Gezahlte Zinsen“ für die Jahre 1985 und 1988 und zeigt, dass diese eng der logarithmischen Verteilung folgen.

Da für den europäischen Raum bisher keine derartigen Untersuchungen vorliegen, wird nun überprüft, ob bei deutschen Steuererklärungen von einer Anpassung zu der Benford-Verteilung ausgegangen werden kann.

4. Steuererklärungen 2001

Nach Schätzungen des Bundesfinanzministeriums gehen bei einem jährlichen Steueraufkommen von > 450 Mrd. Euro allein 75 Mrd. Euro durch Steuerhinterziehung verloren. Hierbei sind strafrechtlich unsanktionierte Formen der Steuervermeidung nicht eingeschlossen.²⁰ Es besteht daher seitens des Bundes und der Länder großes Interesse an einer möglichst umfassenden Aufspürung von manipulierten Steuerangaben. Untersucht wurden die Steuererklärungen eines nordrheinwestfälischen Finanzamtes für das Jahr 2001. Der Datensatz enthält rund 21 Millionen Einträge und ist in verschiedene Sachbereiche (kurz: SB) gegliedert (vgl. Anhang I). Ferner sind die einzelnen SB nach Steuernummern in 2000er und 5000er Bereiche getrennt, wobei letzterer Bereich dann zugeteilt wird, wenn mindestens eine Gewinneinkunftsart vor-

liegt.²¹ Da es sich um Daten *eines* Finanzamtes handelt, sind diese kaum repräsentativ für den Bundesdurchschnitt. Es soll jedoch untersucht werden, inwieweit sich die Ziffernanalyse für die Fälschungsaufspürung eignet, und da Steuerrevisionen sich jeweils auf einzelne Finanzämter bzw. Teilbereiche dieser beziehen, stellt die Repräsentativität hier kein großes Problem dar.

Die Anpassung der ersten, zweiten und ersten beiden Ziffern wird mittels des χ^2 -Tests überprüft, wobei zusätzlich zu der Teststatistik der jeweilige p -Wert in Prozent angegeben wird.²² Die einzelnen SB werden getrennt untersucht und einzelne Teilbereiche werden – dort wo es angebracht scheint – gesondert betrachtet. Die Nullhypothese lautet, dass jeder betrachtete Bereich dem benfordschen Ziffernverhalten folgt. Da keine Informationen seitens der Steuerbehörde über mögliche Fälschungen in dem betrachteten Datensatz vorliegen, kann bei einer Ablehnung von H_0 jedoch nicht unterschieden werden, ob der untersuchte Bereich generell keine logarithmische Ziffernverteilung aufweist, ob die Ablehnung auf die mangelnde Repräsentativität zurückzuführen sind oder ob tatsächlich Fälschungen in signifikanter Weise stattfanden.

Die Ergebnisse beziehen sich, sofern nicht anders angegeben, auf die Zusammenfassung beider Steuernummernbereiche. Detaillierte Werte sowie eine Beschreibung des Vorgehens finden sich im Anhang Cff.

SB 18 Mantelbogen Auf dem Mantelbogen der Steuererklärung im SB 18 befanden sich übermäßig viele Einträge (rund 34%) mit einer Sieben als erster Ziffer. Der Grund hierfür liegt in den Feldern 73 und 74. Hier wird der steuerfrei belassene Arbeitslohn aus sog. geringfügiger Beschäftigung des Ehemannes und der Ehefrau eingetragen. Ein Großteil der Steuererklärenden arbeiten das gesamte Jahr bei einem Lohn von 630DM, was in einem Jahreseintrag von 7560DM resultiert. Nach Entfernung dieser beiden Felder steigt die Übereinstimmung mit Benfords ersten Ziffern auf einen χ^2 -Wert von 14.664 (entspricht: 6.60%). Die Teststatistik der zweiten Ziffern liegt bei 2.069 (99.03%) und die der ersten beiden bei 86.841 (51.49%).

SB 19/29 Ausländische Einkünfte und Steuern Die SB 19 und 29 enthalten die ausländischen Einkünfte und Steuern für den Steuerpflichtigen (SB 19) und den Ehepartner (SB 29). Während die Übereinstimmung der ersten Ziffern mit einer Teststatistik von 8.081 (42.56%) sehr gut war, kann bei den zweiten Ziffern nur ein χ^2 -Wert von 82.887 ($4,306 \cdot 10^{-14}\%$) erreicht werden. Auch der Test der ersten beiden Ziffern muss die Nullhypothese bei einem χ^2 -Wert von 207.815 ($1.085 \cdot 10^{-9}\%$) verwerfen.

SB 22 Einkünfte aus selbständiger Arbeit Mit der Anlage GSE werden die Einkünfte aus selbständiger Arbeit erhoben. Hierbei weisen die ersten Ziffern mit einer Teststatistik von 12.238 (14.09%) eine gute, die zweiten Ziffern mit 19.289 (2.28%) eine noch passable Anpassung an Benfords Gesetz auf. Der gemeinsame Test der ersten beiden Ziffern liegt leicht

über den Ergebnissen der zweiten Ziffern mit einem χ^2 -Wert von 116.236 (2.36%). Fraglich ist, ob die schwache Übereinstimmung der zweiten Ziffern bereits ein Indiz für das Vorliegen von Manipulationen ist. Der SB würde sich hierfür eignen und eine nähere Untersuchung mittels eines Vergleichs gefälschter und unmanipulierter Einträge scheint angebracht.

SB 25 Einkünfte aus Vermietung und Verpachtung Auch bei dieser Anlage V zur Steuererklärung konnte die Nullhypothese bestätigt werden. Im Gegensatz zu vorherigem SB wird die Anpassung bei Betrachtung der zweiten Ziffern jedoch besser. Während bei den ersten Ziffern ein χ^2 -Wert von 15.730 (4.64%) auftritt, ergibt die Betrachtung der zweiten Ziffern einen Wert von 14.743 (9.82%). Der gemeinsame Test der ersten beiden Ziffern liegt mit einer Teststatistik von 114.315 (3.12%) unter diesen Werten. Ebenfalls bei diesem SB sind ungeplante Manipulationen denkbar. Beispielsweise könnte versucht werden, mittels fiktiver Mietforderungen das Steueraufkommen des Ehepartners, von Verwandten oder Bekannten zu schmälern. Daher eignet sich auch dieser Bereich für eingehendere Untersuchungen seitens der Finanzbehörden.

SB 36/37 Kinder Bei der Anlage Kinder kann nicht von einer Benford-Verteilung ausgegangen werden. Grund hierfür sind die Angaben auf dieser Anlage. Neben den Geburtsdaten der Kinder wird die Anzahl der Monate mit vollen bzw. halbem Kinderfreibetrag, Länderangaben, das im Jahr 2000 ausgezahlte Kindergeld und Betreuungsfreibeträge angegeben. Die einzigen Währungsangaben – das Kindergeld und die Betreuungsbeträge – sind jedoch festgelegte Werte. Für das Kindergeld überwiegen beispielsweise Beträge in Höhe von 3240DM.

SB 47/48 Einkünfte aus nichtselbständiger Arbeit A priori kann man in diesem SB nicht von Fälschungsmöglichkeiten ausgehen, da auf der Anlage N lediglich die Angaben der Lohnsteuerkarten übernommen werden. Dass dennoch keine Übereinstimmung mit Benfords Gesetz vorliegt, mag daran liegen, dass die untersuchten Daten lediglich aus einem Finanzamt stammen. Die Bezirkseinteilung der Finanzämter folgt geographischen Gesichtspunkten und ist somit weder für die Stadt – die aus mehreren Finanzämtern bestehen kann – noch für die Region repräsentativ.

Vielmehr ist davon auszugehen, dass die Lage der Steuerbezirke nicht unabhängig von der Höhe des Bruttoarbeitslohns und damit auch nicht unabhängig von den restlichen Angaben dieses Sachbereiches ist, da ein Großteil der Angaben im SB 47/48 aus Werten besteht, die basierend auf dem Bruttoarbeitslohn berechnet werden – beispielsweise die Lohnsteuer, der Solidaritätszuschlag oder die Kirchensteuer. Testet man nur den Bruttoarbeitslohn der ersten Lohnsteuerkarte, so versagt bereits hier die Übereinstimmung der ersten Ziffern. Der Datensatz weist übermäßig viele Fünfen und Sechsen an erster Stelle auf, was bei Abstraktion von der vorliegenden Unwahrscheinlichkeit einer Fälschung für eine Manipulation spräche.

SB 52 Sonderausgaben Der SB eignet sich a priori nur beschränkt für Fälschungen. Er enthält die Beiträge zur Kranken-, Unfall-, Lebens- und Haftpflichtversicherung sowie Spenden für förderungsfähige Zwecke. Von einer Übereinstimmung mit der Benford-Ziffernverteilung kann nicht ausgegangen werden. Auch eine separate Untersuchung der Spenden änderte dieses Ergebnis nicht. Wenn man davon ausgeht, dass Versicherungen und Spenden teilweise proportional zum Einkommen stehen, könnte ein Grund für die Verwerfung der Nullhypothese analog zum SB 47 in der geographischen Lage des Steueramtes liegen.

SB 53 Außergewöhnliche Belastungen Auf der vierten Seite der Einkommenserklärung befindet sich der SB 53. Hier sind vom Steuererklärenden nur die Felder 56 und 57 mit dem Grad der Behinderung Hinterbliebener und Behinderter auszufüllen. Die restlichen Angaben werden von den Finanzbeamten aufgrund gesetzlicher Pauschalbeträge und eingereichter Dokumente eingetragen. Von einer Manipulation ist daher nicht auszugehen, dennoch konnte keine Übereinstimmung im Ziffernverhalten festgestellt werden. Eine Auslassung oben genannter Felder verbessert die Anpassung nur leicht.

SB 54 Einkünfte aus Kapitalvermögen Dieser SB der Anlage GSE enthält die Angaben zu gezahlten und erhaltenen Zinsen und Dividenden. Nach der Untersuchung von Nigrini (1996) ist zu vermuten, dass der analoge SB amerikanischer Steuererklärungen für Fälschungen genutzt werden kann. Der SB zeigt hier ein uneinheitliches Verhalten. Während bei den 5000er Steuernummern bei einem χ^2 -Wert von 13.189 (10.55%) für die ersten Ziffern von einer guten Anpassung ausgegangen werden kann, weist die größere Klasse der 2000er Steuernummern nur ein χ^2 -Wert von 27.216 (0.064%) auf. Bezüglich der zweiten Ziffern lehnt der Test in beiden Fällen die Nullhypothese signifikant ab. Es ist jedoch zu bemerken, dass in beiden Fällen übermäßig viele Fünfen und Sechsen an erster Stelle auftreten, was nach den obigen Hypothesen ein Indiz für ungeplante Manipulation ist. Ein separater Test der erhaltenen und gezahlten Zinsen ändert das Ergebnis kaum. Im Gegensatz zu Nigrinis Vermutung, dass die erhaltenen Zinsen nach unten manipuliert werden, weist dieses Feld zu wenige Einsen und übermäßig viele Fünfen und Sieben als erste Ziffern auf.

SB 55 Sonstige Einkünfte Mit der Anlage SO werden sonstige Einkünfte wie Leibrenten, Unterhaltsleistungen oder private Veräußerungsgeschäfte erfasst. Eine Steuervermeidung wäre durch Verkleinerung der Beträge möglich, weshalb übermäßig viele kleine erste Ziffern ein Indiz für Manipulationen wären. Der Zifferntest der ersten Ziffern weist zwar weniger Einsen (28.50%) , jedoch mehr Zweien (22.11%) und Dreien (18.98%) als erwartet auf. Der χ^2 -Test von 619.61 lehnte die Nullhypothese bezüglich der ersten Ziffern ab und verwarf bei den zweiten Ziffern H_0 nur knapp mit 19.57 (2.08%). Testet man lediglich die zweite Seite dieser Anlage, die private Veräußerungsgeschäfte enthält, so ergibt sich eine gute Anpassung der ersten Ziffern bei einem χ^2 -Wert von 5.69 (68.17%), als auch der zweiten Ziffern bei

einem χ^2 -Wert von 10.19 (33.50%). Der Test der ersten beiden Ziffern bestätigt erwartungsgemäß dieses Ergebnis mit einer Teststatistik von 89.73 (42.87%). Es kann also hinsichtlich der zweiten Seite von einer Übereinstimmung im Ziffernverhalten ausgegangen werden. Das Abweichen der ersten Seite der Anlage könnte bezüglich der Renteneinträge an der Proportionalität der Rente zum Einkommen liegen und damit analog zur Argumentation des SB 47 zu sehen sein.

SB 87 Werbungskosten Der letzte SB 87 der Anlage N enthält die Werbungskosten bei Einnahmen aus nichtselbständiger Arbeit und weist keine Übereinstimmung des Ziffernverhaltens auf. Auf diesem Formulareil werden neben den Fahrten zwischen Wohn- und Arbeitsstätte, Beiträgen zu Berufsverbänden auch Aufwendungen für Arbeitsmittel und Werbungs- sowie Fortbildungskosten veranschlagt. All diese Einträge weisen ein gesetzliches Maximum auf, so dass zwar von einer manipulativen Erhöhung der Beträge ausgegangen werden kann, die Fälschungsmöglichkeiten jedoch begrenzt bleiben und die Abweichung vom benfordschen Ziffernverhalten auf die Existenz der Pauschalbeträge zurückzuführen ist.

E. Diskussion

Bezüglich der empirischen Anwendungsmöglichkeiten scheinen dem Phänomen der ersten Ziffern keine Grenzen gesetzt zu sein. Neben dem aktuellen Einsatz in der Bilanzprüfung wurde gezeigt, dass auch die Wirtschaftswissenschaften sich die logarithmische Verteilung zu Nutze machen können. Besonders die Entscheidung über die Qualität von Zeitreihendaten könnte mittels des Benford-Tests auf Natürlichkeit bzw. Unverzerrtheit ermöglicht werden.

Auf dem Anwendungsgebiet der Steuerrevision sind indes noch zu viele Fragen offen, um von einer derzeitigen praktischen Nutzung ausgehen zu können. Es sind hier noch Untersuchungen mit der sicheren Kenntnis des möglichen Fälschungsgrades von Nöten, um zwischen einer zufälligen – oder auf der Datenauswahl basierenden – Abweichung und einer tatsächlichen Fälschung zu unterscheiden. Ferner ist fraglich, ob einzelne Steuererklärungen mit entsprechender Genauigkeit hierbei ausgewählt werden können.

Eng damit verbunden ist die Konstruktion geeigneter Anpassungstests. Der klassische χ^2 -Test hat sich zwar als brauchbar erwiesen, es stehen mit den neuen Invarianz-Tests jedoch Verfahren zur Verfügung, die spezifisch auf den Eigenschaften der Benford-Verteilung aufbauen. Die Transformationsinvarianzen scheinen – in einer Zeit computergestützter Verfahren – für die Ziffernanalyse durchaus geeignet zu sein, zumal die Zusammenfassung der Testreihe in der Maßzahl der Varianz eine einfache Interpretation ermöglicht. Eine nähere theoretische Betrachtung dieser Verfahren ist aber angebracht, bevor die praktischen Einsatzgebiete sondiert werden können. Auch die Quantifizierung der erforderlichen Anpassungsgüte müsste sowohl theoretisch als auch empirisch fundiert werden, wobei noch fraglich ist, ob theoreti-

sche Verwerfungsbereiche gefunden werden können, die auch robust gegen zufällige Ausreißer auf einzelnen Transformationsstufen sind. Erst nach eingehenden Untersuchungen kann sich herausstellen, ob in großen Datensätzen zwischen einem „Rauschen“ und tatsächlicher Verzerrung unterschieden werden kann. Neuere Ansätze basieren auf neuronalen Netzwerken und erzielen erstaunliche Genauigkeit bei der Fälschungsaufspürung (vgl. Busta und Weinberg (1998)).

Anmerkungen

¹Die Wettidee ist inspiriert von Weaver (1963).

²Vor der Verbreitung von Taschenrechnern und Computern benutzte man die Logarithmentafeln, um Multiplikationen großer Zahlen mittels der Addition ihrer logarithmischen Mantissen zu bewerkstelligen.

³Diese deduktive Herleitung ist antichronologisch im Bezug auf die induktive Entwicklung von Benfords Gesetz zu sehen.

⁴Das Korollar folgt direkt aus Satz B.5. Vgl. auch Schatte (1998), S. 391.

⁵Für weitere numerische Werte vgl. Tabelle 2.

⁶Die Herleitung des Satzes stammt von K. Schürger und ist bisher unveröffentlicht.

⁷Die σ -Algebra wird äquivalent durch die Mantissenfunktion generiert. Zum Zusammenhang zwischen der Ziffernfunktion $D_i^{(b)}$ und der Mantissenfunktion M_b sei an die Definitionen B.1 und B.2 erinnert.

⁸Eine statistische Herleitung scheiterte bisher an der Tatsache, dass kein skaleninvariantes W.-Maß auf den Boreluntermengen von \mathbb{R}^+ existiert. Wenn es ein solches gäbe, wäre das Maß des Intervalls $[0, 1[$ gleich dem Maß jedes Intervalls $[0, \alpha[$, welches wegen der σ -Additivität den Wert Null haben müsste (Hill 1995b, S. 324).

⁹Der Beweis basiert auf der Zuordnung der Skaleninvarianz zu der Invarianz unter irrationalen Rotationen $x \rightarrow (x + s) \pmod{1}$ auf dem Kreis. Der interessierte Leser sei auf (Hill 1996, S. 358) verwiesen. Der Beweis zeigt auch, dass obige Definition B.11 ein wenig zu stark ist und durch die Hypothese der Skaleninvarianz bezüglich einer beliebigen positiven Zahl – die nicht von der Form b^r für ein rationales $r > 0$ ist – ersetzt werden kann.

¹⁰Da in Definition B.13 nur die Invarianz bezüglich Exponenten der Ausgangsbasis gefordert wird, mag diese Begriffslegung auf den ersten Blick zu schwach wirken. Der Beweis zu folgendem Satz B.14 zeigt, dass dies für die Baseninvarianz bezüglich *beliebiger* Basen ausreicht.

¹¹Für alle $S \in \mathcal{M}_b$ gilt: $\delta_1(S) = 1$, falls $S_1 \subseteq S$ und $\delta_1(S) = 0$, sonst.

¹²Wie der Beweis zeigt, korrespondiert die Baseninvarianz zur Invarianz unter Multiplikationen $x \rightarrow nx \pmod{1}$ (Hill 1996, S. 368).

¹³Die empirischen Schranken basieren auf den Erfahrungen Nigrinis und beziehen sich auf Unternehmungsbarauszahlungen mit weniger als einer Millionen Daten.

¹⁴Nigrini benutzt die folgende Spezifikation für sein Verzerrfaktormodell. Er betrachtet eine geometrische Folge mit $q = 10^{1/N}$ und $a_1 = 10$.

¹⁵Die Eigenschaft der Summeninvarianz eignet sich hingegen weniger für ein Testverfahren, da hierzu nicht nur die Mantissen berechnet, sondern diese hiernach sortiert und gruppenweise untersucht werden müssen.

¹⁶Ferner ist zu beachten, dass ein Datensatz der mittels einer geometrischen Folge generiert wurde, keine unabhängigen Einträge enthält und folglich nicht mittels des χ^2 -Tests auf die Anpassung überprüft werden kann.

¹⁷Obwohl eine Abhängigkeit bei der Untersuchung der Inflationsraten nicht völlig ausgeschlossen werden kann, ist eine Anwendung des χ^2 -Test insofern möglich, da lediglich ein Vergleich der Teststatistiken einzelner Transformationsstufen stattfindet und folglich kein kritischer Wert benutzt wird. In diesem Fall wird der Wert der χ^2 -Statistik als Abstandsmaß benutzt.

¹⁸Es gibt jedoch keine Konstante k , so dass ein Nicht-Benford Datensatz mittels Multiplikation mit k zu einem Benford Datensatz wird (Raimi 1969, S. 118).

¹⁹Der Datensatz des Tax Compliance Measurement Program (TCMP) ist anonymisiert und von Finanzbeamten auf Fälschungen überprüft. Dubiose Einträge werden nicht markiert, so dass die Benutzer des Datensatzes keine Anhaltspunkte für die Unterscheidung zwischen manipulierten und „natürlichen“ Daten haben.

²⁰Bundesministerium der Finanzen: Monatsbericht Dezember 2002, S. 57.

²¹Gewinneinkunftsarten sind hierbei Einkünfte aus Land- und Forstwirtschaft, aus Gewerbebetrieben oder selbständiger Arbeit. Die Bereichseinteilung bezieht sich auf die jeweils aktuelle Steuererklärung, es können daher auch in 2000er Bereichen Gewinneinkünfte vorliegen.

²²Die folgenden Untersuchungen der Steuererklärungen werden mittels des unter <http://benford.verita.de> verfügbaren Programmes BDA - Benford Digital Analysis durchgeführt.

Literatur

Bamberg und Baur 1998 BAMBERG, Günter ; BAUR, Franz: *Statistik*. 10. überarbeitete Auflage. Oldenbourg, 1998

Benford 1938 BENFORD, Frank: The Law of anomalous Numbers. In: *Proceedings of the American Philosophical Society* 78 (1938), Nr. 4, S. 551–572

Bronstein u. a. 1999 BRONSTEIN, I.N. ; SEMENDJAJEW, K.A. ; MUSIOL, G. ; MÜHLIG, H.: *Taschenbuch der Mathematik*. 4., überarb. und erw. Auflage. Harri Deutsch, 1999

Büning und Trenkler 1978 BÜNING, Herbert ; TRENKLER, Götz: *Nichtparametrische statistische Methoden*. de Gruyter, 1978

Busta und Sundheim 1992 BUSTA, Bruce ; SUNDHEIM, Richard: *Detecting manipulated tax returns with the use of Benford's Law*. April 1992. – Working Paper St. Cloud State University

Busta und Weinberg 1998 BUSTA, Bruce ; WEINBERG, Randy: *Using Benford's Law and Neural Networks as a Review Procedure*. January 1998. – St. Cloud State University, Minnesota. Department of Business.

Feldstein und Goodman 1976 FELDSTEIN, Alan ; GOODMAN, Richard: Convergence Estimates for the Distributions of Trailing Digits. In: *Journal of the Association for Computing Machinery* 23 (1976), April, S. 287–297

Gibbons 1985 GIBBONS, Jean D.: *Nonparametric Statistical Inference*. 2. überarbeitete und erweiterte Auflage. Marcel Dekker, 1985

- Hill 1988** HILL, Theodore P.: Random-Number Guessing and the First Digit Phenomenon. In: *Psychological Reports* 62 (1988), S. 967–971
- Hill 1995a** HILL, Theodore P.: Base-Invariance Implies Benford’s Law. In: *Proceedings of the American Mathematical Society* 123 (1995), S. 887–895
- Hill 1995b** HILL, Theodore P.: The Significant-Digit Phenomenon. In: *The American Mathematical Monthly* 102 (1995), S. 322–327
- Hill 1996** HILL, Theodore P.: A Statistical Derivation of the Significant-Digit Law. In: *Statistical Science* 10 (1996), Nr. 4, S. 354–363
- Hill 1999** HILL, Theodore P.: The difficulty of faking data. In: *Chance* 26 (1999), S. 8–13
- Jamain 2001** JAMAIN, Adrien: *Benford’s Law*. 2001. – Unpublished Dissertation Report, Department of Mathematics, Imperial College, London.
- Ley 1996** LEY, Eduardo: On the Peculiar Distribution of the U.S. Stock Indices First Digits. In: *The American Statistician* 50 (1996), Nr. 4, S. 311–314
- Newcomb 1881** NEWCOMB, Simon: Note on the Frequency of Use of the Different Digits in Natural Numbers. In: *Amer. J. Math.* 4 (1881), S. 39–40
- Nigrini 1996** NIGRINI, Mark J.: A taxpayer compliance application of Benford’s law. In: *Journal of the American Taxation Association* 18 (1996), S. 72–91
- Nigrini 2000** NIGRINI, Mark J.: *Digital Analysis using Benford’s Law*. Global Audit, 2000
- Pinkham 1961** PINKHAM, Roger S.: On the distribution of first significant digits. In: *Annals of Mathematical Statistics* 32 (1961), S. 1223–1230
- Raimi 1969** RAIMI, R.A.: The peculiar distribution of first digits. In: *Scientific American* 221 (1969), S. 109–120
- Schatte 1998** SCHATTE, Peter: On Benford’s Law to Variable Base. In: *Statistics & Probability Letters* 37 (1998), S. 391–397
- Weaver 1963** WEAVER, Warren: *Lady Luck. The Theory of Probability*. Dover Publications, 1963

Anhang

A. Tabellen

| d | $P(D_1 = d)$ | $P(D_2 = d)$ | $P(D_3 = d)$ | $P(D_4 = d)$ |
|---|--------------|--------------|--------------|--------------|
| 0 | - | .11967926840 | .10178436490 | .10017614540 |
| 1 | .30102999570 | .11389010390 | .10137597840 | .10013689220 |
| 2 | .17609125910 | .10882149850 | .10097219580 | .10009767100 |
| 3 | .12493873650 | .10432956040 | .10057293260 | .10005849910 |
| 4 | .09691001300 | .10030820300 | .10017808800 | .10001937270 |
| 5 | .07918124605 | .09667723578 | .09978757705 | .09998028723 |
| 6 | .06694678975 | .09337473577 | .09940130907 | .09994123464 |
| 7 | .05799194701 | .09035198920 | .09901920603 | .09990223654 |
| 8 | .05115252247 | .08757005334 | .09864118340 | .09986328509 |
| 9 | .04575749054 | .08499735264 | .09826716225 | .09982437469 |

Tabelle 1: Relative Häufigkeit der Ziffer d an erster bis vierter Stelle.

Generiert mittels $P(D_n = d) = \sum_{k=10^{n-1}}^{10^n-1} \log(1 + \frac{1}{10k+d})$, gerundet auf zehn signifikante Stellen.

| $\downarrow D_1, D_2 \rightarrow$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | - |
|-----------------------------------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|
| 1 | .0413926 | .0377888 | .0347621 | .0321834 | .0299637 | .0280300 | .0263289 | .0248221 | .0234829 | .0222757 | .30102999 |
| 2 | .0211893 | .0202037 | .0193032 | .0184841 | .0177301 | .0170333 | .0163897 | .0157955 | .0152381 | .0147221 | .17609126 |
| 3 | .0142390 | .0137891 | .0133640 | .0129637 | .0125883 | .0122338 | .0119001 | .0115831 | .0112828 | .0109949 | .12493873 |
| 4 | .0107238 | .0104653 | .0102194 | .00998597 | .00976096 | .00954434 | .00934039 | .00914479 | .00895340 | .00877468 | .096910012 |
| 5 | .00860014 | .00843407 | .00827216 | .00811877 | .00796955 | .00782454 | .00768804 | .00755145 | .00742342 | .00729961 | .079181247 |
| 6 | .00718000 | .00706035 | .00694926 | .00683813 | .00673125 | .00662858 | .00653021 | .00643606 | .00634195 | .00624775 | .066946803 |
| 7 | .00616215 | .00607221 | .00599082 | .00590943 | .00582805 | .00575087 | .00567804 | .00560517 | .00553225 | .00546363 | .057991933 |
| 8 | .00539501 | .00533069 | .00526633 | .00520197 | .00513761 | .00507750 | .00502169 | .00496159 | .00490578 | .00485423 | .051152524 |
| 9 | .00479838 | .00474683 | .00469528 | .00464373 | .00459648 | .00454918 | .00450193 | .00445464 | .00440734 | .00436435 | .045757488 |
| - | .11967926 | .11389004 | .10882147 | .10432955 | .10030826 | .096677189 | .093374706 | .090352004 | .087570073 | .084997425 | |

Tabelle 2: Die gemeinsame Verteilung nach Benford

Die gemeinsame Verteilung wurde auf sechs signifikante Nachkommastellen, die Randverteilung auf acht signifikante Nachkommastellen gerundet.

| Test / α | 0.05 | 0.01 |
|-----------------|-------------|-------------|
| χ_8^2 | 15.50731249 | 20.09015916 |
| χ_9^2 | 16.91896016 | 21.66604759 |
| χ_{89}^2 | 112.0219553 | 122.9421709 |
| χ_{99}^2 | 123.2252305 | 134.6414896 |
| χ_{899}^2 | 969.8647614 | 1000.574515 |

Tabelle 3: Kritische Werte des χ^2 -Tests.

Die kritischen Wert des KS-Tests können für den zweiseitigen Test bei mehr als 35 Beobachtungen approximiert werden durch: $\frac{z_\alpha}{\sqrt{n}}$, wobei $z_{0.05} = 1.36$ und $z_{0.01} = 1.63$ ist (Gibbons 1985, S. 81).

B. Testverfahren

| | MAD | Anpassungsgüte |
|----------------------|----------------------------------------|-----------------------|
| Erste Ziffern | $[0, 4 \cdot 10^{-3}[$ | eng |
| | $[4 \cdot 10^{-3}, 8 \cdot 10^{-3}[$ | akzeptabel |
| | $[8 \cdot 10^{-3}, 12 \cdot 10^{-2}[$ | marginal akzeptabel |
| | $\geq 12 \cdot 10^{-2}$ | keine Übereinstimmung |
| Zweite Ziffern | $[0, 8 \cdot 10^{-3}[$ | eng |
| | $[8 \cdot 10^{-3}, 12 \cdot 10^{-3}[$ | akzeptabel |
| | $[12 \cdot 10^{-3}, 16 \cdot 10^{-2}[$ | marginal akzeptabel |
| | $\geq 16 \cdot 10^{-2}$ | keine Übereinstimmung |
| Erste beiden Ziffern | $[0, 6 \cdot 10^{-4}[$ | eng |
| | $[6 \cdot 10^{-4}, 12 \cdot 10^{-3}[$ | akzeptabel |
| | $[12 \cdot 10^{-3}, 18 \cdot 10^{-3}[$ | marginal akzeptabel |
| | $\geq 18 \cdot 10^{-3}$ | keine Übereinstimmung |

Tabelle 4: Anpassungsschranken der MAD nach Nigrini (2000).

I. Spezifische Benford Tests

1. Baseninvarianz

| Basis: | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 |
|----------|---------|----------|---------|---------|---------|---------|---------|---------|
| MAD | 0.02280 | 0.02524 | 0.02011 | 0.03084 | 0.03166 | 0.04904 | 0.04860 | 0.00272 |
| MSE | 0.00074 | 0.00084 | 0.00058 | 0.00189 | 0.00160 | 0.00338 | 0.00296 | 0.00001 |
| χ^2 | 9.20709 | 12.34935 | 4.76277 | 7.20958 | 6.54025 | 8.92398 | 4.04719 | 0.00497 |

Tabelle 5: Test auf Baseninvarianz

2. Skaleninvarianz

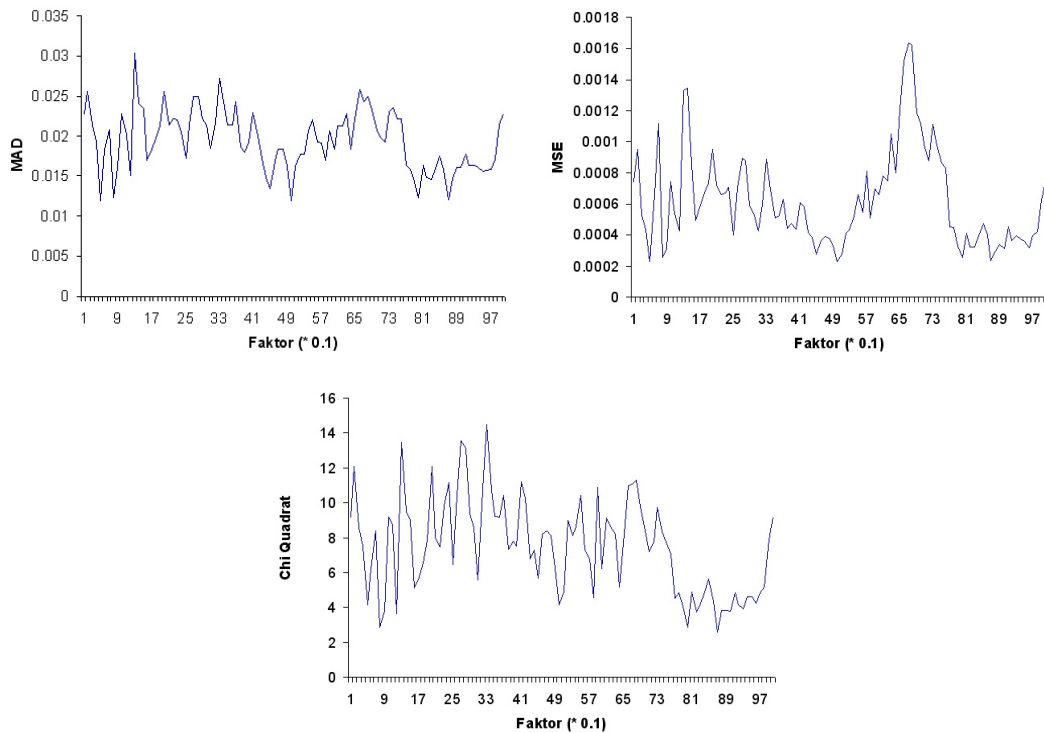


Abbildung 1: Skaleninvarianz-Test des Datensatzes CPI.

C. Steuererklärungen

I. Dokumentation der Bearbeitung

In jedem SB wurden Kennzahlenfelder, wie z.B. Länderschlüssel, Indikatorfelder oder Steuerklassennummern entfernt. Nach einem ersten χ^2 -Test wurden die schwach übereinstimmenden SB erneut auf Felder überprüft, die ggf. durch feste gesetzliche Vorgaben die unterstellte

Natürlichkeit des Datensatzes beeinflussen. Die Resultate beziehen sich wiederum auf absolute Werte.

| Datei | Zeilen | Spalten | Felder |
|-------------|---------|---------|------------|
| 2er-SB14-25 | 6561 | 198 | 1 299 078 |
| 2er-SB36-47 | 27705 | 178 | 4 931 490 |
| 2er-SB48-51 | 7462 | 62 | 462 644 |
| 2er-SB52-54 | 28334 | 172 | 4 873 448 |
| 2er-SB55-88 | 28471 | 206 | 5 865 026 |
| 5er-SB14-25 | 3706 | 224 | 830 144 |
| 5er-SB36-47 | 4920 | 146 | 718 320 |
| 5er-SB48-51 | 1918 | 94 | 180 292 |
| 5er-SB52-54 | 6257 | 168 | 1 051 176 |
| 5er-SB55-88 | 5353 | 184 | 984 952 |
| Summe | 120 687 | 1 634 | 21 196 570 |

Tabelle 6: Aufbau der Rohdaten der Steuererklärungen.

| SB | Beschreibung | Vordruck | Anmerkung |
|-------|------------------------------------------------------------|---------------|-----------|
| 14 | Getrennte Veranlagung | EST 5 A | |
| 18 | Mantelbogen unbeschränkt Steuerpflichtige | | |
| 19/29 | Ausländische Einkünfte und Steuern | Anlage AUS | ab 1990 |
| 22 | Einkünfte aus selbstständiger Arbeit | Anlage GSE | |
| 25 | Einkünfte aus Vermietung und Verpachtung | Anlage V | |
| 36/37 | Angabe zu Kindern | Anlage Kinder | ab 1996 |
| 47/48 | Einkünfte aus nichtselbständiger Arbeit | Anlage N | ab 1990 |
| 52 | Sonderausgaben | | ab 1999 |
| 53 | Außergewöhnliche Belastungen | | ab 1999 |
| 54 | Einkünfte aus Kapitalvermögen | Anlage KAP | ab 2000 |
| 55 | Sonstige Einkünfte | Anlage SO | ab 2000 |
| 87/88 | Werbungskosten bei Einnahmen aus nichtselbständiger Arbeit | Anlage N | |

Tabelle 7: Auflistung der Sachbereiche.

| SB | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 14 | 0.21324 | 0.10662 | 0.10478 | 0.13419 | 0.15625 | 0.10478 | 0.06801 | 0.04596 | 0.06618 |
| 18 | 0.10215 | 0.08764 | 0.09066 | 0.08613 | 0.11151 | 0.16470 | 0.34089 | 0.00937 | 0.00695 |
| 18 angepasst | 0.32484 | 0.24204 | 0.08280 | 0.06369 | 0.05732 | 0.09554 | 0.07643 | 0.02548 | 0.03185 |
| 19 | 0.29796 | 0.19242 | 0.12620 | 0.09168 | 0.07612 | 0.06480 | 0.05546 | 0.05207 | 0.04329 |
| 22 | 0.30414 | 0.20134 | 0.12713 | 0.08090 | 0.07968 | 0.06083 | 0.05414 | 0.04927 | 0.04258 |
| 25 | 0.29940 | 0.16730 | 0.11836 | 0.10107 | 0.08328 | 0.07261 | 0.05826 | 0.05053 | 0.04918 |
| 36 | 0.19706 | 0.02230 | 0.74778 | 0.02123 | 0.00276 | 0.00138 | 0.00020 | 0.00571 | 0.00158 |
| 47 2000er | 0.25756 | 0.13231 | 0.11470 | 0.10251 | 0.10520 | 0.10211 | 0.07622 | 0.05931 | 0.05007 |
| 47 5000er | 0.22317 | 0.11463 | 0.09382 | 0.07958 | 0.27519 | 0.06923 | 0.05323 | 0.04797 | 0.04319 |
| 52 | 0.38595 | 0.11821 | 0.09292 | 0.07850 | 0.08166 | 0.07937 | 0.05785 | 0.05364 | 0.05189 |
| 53 | 0.30035 | 0.20547 | 0.12149 | 0.10952 | 0.07212 | 0.05994 | 0.04199 | 0.04765 | 0.04146 |
| 54 2000er | 0.29107 | 0.17054 | 0.12550 | 0.09937 | 0.08502 | 0.07052 | 0.05984 | 0.05324 | 0.04489 |
| 54 5000er | 0.30090 | 0.17558 | 0.12473 | 0.09458 | 0.08254 | 0.07242 | 0.05789 | 0.04813 | 0.04322 |
| 55 | 0.28504 | 0.22113 | 0.18975 | 0.09188 | 0.05393 | 0.05111 | 0.04007 | 0.03631 | 0.03078 |
| 87 | 0.20526 | 0.30576 | 0.21184 | 0.08853 | 0.06855 | 0.04399 | 0.03105 | 0.02554 | 0.01947 |
| Mittel | 0.26587 | 0.16422 | 0.16483 | 0.08823 | 0.09274 | 0.07422 | 0.07144 | 0.04068 | 0.03777 |
| Benford | 0.30103 | 0.17609 | 0.12494 | 0.09691 | 0.07918 | 0.06695 | 0.05799 | 0.05115 | 0.04576 |

Tabelle 8: Relative Häufigkeit der ersten Ziffern nach SB.

| SB | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 14 | 0.13548 | 0.10538 | 0.12903 | 0.09892 | 0.10323 | 0.09462 | 0.07527 | 0.07957 | 0.09677 | 0.08172 |
| 18 | 0.07881 | 0.07482 | 0.09261 | 0.09046 | 0.10120 | 0.26924 | 0.06838 | 0.06440 | 0.08586 | 0.07421 |
| 18 angepasst | 0.11613 | 0.12258 | 0.11613 | 0.11613 | 0.11613 | 0.07097 | 0.09677 | 0.08387 | 0.08387 | 0.07742 |
| 19 | 0.06705 | 0.12650 | 0.11707 | 0.11852 | 0.10149 | 0.10584 | 0.10330 | 0.09641 | 0.08046 | 0.08336 |
| 22 | 0.09241 | 0.10759 | 0.11456 | 0.10696 | 0.10316 | 0.11392 | 0.09051 | 0.08354 | 0.09177 | 0.09557 |
| 25 | 0.11624 | 0.11888 | 0.10520 | 0.09704 | 0.10319 | 0.09264 | 0.09202 | 0.09641 | 0.08938 | 0.08900 |
| 36 | 0.00648 | 0.01343 | 0.69400 | 0.00776 | 0.00939 | 0.00158 | 0.23498 | 0.00888 | 0.02047 | 0.00301 |
| 47 | 0.11998 | 0.11334 | 0.11429 | 0.10310 | 0.10489 | 0.09408 | 0.09122 | 0.08720 | 0.08738 | 0.08452 |
| 52 | 0.10524 | 0.12014 | 0.12816 | 0.12619 | 0.10293 | 0.09412 | 0.09016 | 0.07486 | 0.08718 | 0.07101 |
| 53 | 0.07447 | 0.08030 | 0.17706 | 0.07315 | 0.12628 | 0.07814 | 0.08764 | 0.07569 | 0.17489 | 0.05237 |
| 54 2000er | 0.09975 | 0.11397 | 0.11234 | 0.10736 | 0.10342 | 0.10358 | 0.09734 | 0.08801 | 0.08502 | 0.08921 |
| 54 5000er | 0.10468 | 0.11730 | 0.11347 | 0.10789 | 0.09924 | 0.10016 | 0.09327 | 0.09075 | 0.08830 | 0.08493 |
| 55 | 0.11115 | 0.10830 | 0.11020 | 0.10509 | 0.09831 | 0.09522 | 0.09451 | 0.08987 | 0.09689 | 0.09047 |
| 87 | 0.05656 | 0.09809 | 0.11952 | 0.15459 | 0.08786 | 0.13439 | 0.09114 | 0.07640 | 0.10591 | 0.07554 |
| 88 | 0.05914 | 0.09859 | 0.14856 | 0.15982 | 0.08218 | 0.13856 | 0.07965 | 0.06317 | 0.10098 | 0.06936 |
| Mittel | 0.08957 | 0.10128 | 0.15948 | 0.10487 | 0.09619 | 0.10581 | 0.09908 | 0.07727 | 0.09168 | 0.07478 |
| Benford | 0.11968 | 0.11389 | 0.10882 | 0.10433 | 0.10031 | 0.09668 | 0.09337 | 0.09035 | 0.08757 | 0.08500 |

Tabelle 9: Relative Häufigkeit der zweiten Ziffern nach SB.

II. Änderungen der Sachbereiche

SB 18

Gelöscht wurde Feld 75 Fünftel-Regelung Indikatorfeld.

SB 19

Gelöscht wurden folgende Felder:

| Kennz. | Beschreibung |
|-------------|---------------|
| 18-28 | Indikatorfeld |
| 31,34,37,39 | Staatnummer |
| 42,44,45 | Indikatorfeld |
| 71-82 | Indikatorfeld |

SB 22

Keine Änderungen.

SB 25

Keine Änderungen.

SB 36/37 - Kinder

Gelöscht wurden folgende Felder:

| Kennz. | Beschreibung |
|--------|--------------------------------|
| x0-x4 | Monatsangabe KFB |
| x5 | Ausgezahltes Kindergeld |
| x6 | Geburtsdatum des x.ten Kindes |
| x9 | Betreuungsfreibetrag Indikator |
| 65 | Ausbildungsfreibeträge |
| x | x=1,2,3,4 |

SB 47/48 Einkünfte aus nichtselbständiger Arbeit

47 bezieht sich auf den Steuerpflichtigen, 48 auf den Ehepartner.

Gelöscht wurden folgende Felder:

| Kennz. | Beschreibung |
|--------|-----------------------|
| 67 | Arbeitgeber FA Nummer |
| 68 | Steuerklasse |
| 85 | Veranlagungsgrund |

SB 52

Gelöscht wurden die Kennzahlen 23-26 Indikatorfelder.

SB 53

Gelöscht wurden folgende Felder:

| Kennz. | Beschreibung |
|--------|----------------------------|
| 50 | Summe der Unterhaltsmonate |
| 56/57 | Grad der Behinderung |
| 58 | Hinterbliebenen Anzahl |
| 80 | Länderschlüssel |

SB 54

keine Änderungen.

SB 55

Gelöscht wurden:

40/41 Renteneinkommenindikator.

52/53/56/57 Ertragsanteil der Rente in Prozent.

Die Löschung dieser Felder änderte das Ergebnis qualitativ nicht!

SB 87/88

Gelöscht wurden folgende Felder:

| Kennz. | Beschreibung |
|-------------------|------------------------------------|
| 15 | Bemessungsgrundlage Indikatorfeld |
| 40/41/43/44/46/47 | Tage und Entfernungen in km |
| 35 | Versorgungspauschale Indikatorfeld |
| 61-63 | Schlüsselzahl |

Anmerkung: Feld 52 (Aufwendungen für Arbeitsmittel) weist viele Einträge in Höhe von 200DM auf.

Summary

This paper gives an introduction to the new methodology of Digital Analysis. The basic principles are introduced both theoretically and practically. Furthermore classical statistical goodness-of-fit tests are discussed with respect to their application in the field of digital analysis. A new class of tests, the invariance-test-procedures, are introduced. These tests are based on the properties of the discussed digit distribution. Finally German tax declarations are examined using the introduced methodology. The procedures already used by the German Federal Ministry of Finance to detect fraudulent tax data seem to be, in a statistical sense, too weak to capture the relevant information.

Zusammenfassung

Die vorliegende Arbeit führt in die neue Methodik der Ziffernanalyse („Digital Analysis“) ein. Es werden die theoretischen Grundlagen auf aktuellem Forschungsniveau vorgestellt und diskutiert. Hiernach werden klassische statistische Verfahren der Anpassungsgüte in Hinblick auf den Einsatz in der Ziffernanalyse diskutiert. Ferner wird eine neue Klasse statistischer Tests, die sog. Invarianztests, eingeführt, die speziell auf die Bedürfnisse der Ziffernanalyse abstellen. Letztlich werden deutsche Steuererklärungen mit den vorgestellten Verfahren untersucht. Es ist hierbei davon auszugehen, daß die z.Zt. vom Bundesfinanzministerium eingesetzte Ziffernanalyse stark anfällig bezüglich Fehlspezifikationen ist. Dieses Problem könnte durch den Einsatz der neuen Testverfahren reduziert werden.