International Journal of Pure and Applied Mathematics
Volume 11 No. 1 2004, 39-46

INTEGER POWERS AND BENFORD'S LAW

Werner Hürlimann
Schönholzweg 24
CH-8409 Winterthur
SWITZERLAND
e-mail: whurlimann@bluewin.ch
url: www.geocities.com/hurlimann53

Abstract: The exact probability distribution of the first digit of integer powers up to an arbitrary but fixed number of digits is derived. Based on its asymptotic distribution, it is shown that it approaches Benford's law very closely for sufficiently high powers.

AMS Subject Classification: 11B83, 11K31, 11Y55, 62E15, 62E20
Key Words: integer sequence, powers, first digit, Benford's law, asymptotic distribution

*

It is part of the common knowledge that the first digits of many numerical data sets are not equally distributed. Newcomb [19] and Benford [4] observed that for any extensive collection of real numbers expressed in decimal form, the probability that the first digit j equals d is given by $P(j = d) = \log\left(1 + d^{-1}\right)$, $d \in \{1, ..., 9\}$.

Mathematical explanations of this law have been proposed by Pinkham [23], Hill [11]-[15], Allart [1], Janvresse and de la Rue [16]. In recent years an upsurge of applications of Benford's law have appeared including work by Becker [2], Burke and Kincanon [7], Buck et al [6], Ley [18], Nigrini [20], [21], Nigrini and Mittermaier [22], Vogt [27], Tolle et al [26] and Berger et al [3].

Received: October 15, 2003

Benford's law for integer and other sequences of real numbers has been analyzed extensively by many authors including Brown and Duncan [5], Whitney [29], Diaconis [9], Cohen and Katz [8], Schatte [24] and Gottwald [10]. It is well-known that the sequences of integer powers $\{n^c\}$, $\quad n = 1, 2, \ldots$, for any fixed real number c do not follow Benford's law (e.g. Diaconis [9]). In the present note, we derive the exact probability distribution of the first digit of integer powers up to an arbitrary but fixed number of digits. We obtain its asymptotic distribution and show that it approaches Benford's law very closely for sufficiently high powers.

The following *ceiling function* will be used throughout. For a real number x, let $\lceil x \rceil$ denote the least integer greater than or equal to x. Let us consider the following *m-th power counting sequences*.

Definition 1. For $i \in \{1, \ldots, 9\}$, $\quad m \geq 2$, the element numbered $n \geq 1$ of the sequence $\{a_n(i,m)\}$ denotes the largest integer whose m-th power has n digits and first digit i, provided it exists. If this integer does not exist, the symbol denotes either the largest integer preceding the integer whose m-th power has n digits and smallest first digit greater than i or the largest integer preceding the smallest integer whose m-th power has $n+1$ digits.

Example 1. One has $a_3(1,3) = 5$ because $5^3 = 125 < 6^3 = 216$, which shows that 125 is the largest cube with 3 digits and first digit 1. One has $a_3(4,3) = 7$ because the cube $7^3 = 343$ precedes the cube $8^3 = 512$ with 3 digits and smallest first digit greater than 4. One has $a_2(7,3) = 4$ because $4^3 = 64 < 5^3 = 125$, which shows that the cube 64 precedes the smallest cube 125 with 3 digits.

Lemma 1. *The m-th power counting sequence satisfies the explicit formula*

$$a_n(i,m) = \left\lceil (1+i)^{\frac{1}{m}} \cdot 10^{\frac{n-1}{m}} \right\rceil - 1, \quad n \in N_+. \tag{1}$$

Proof. From the inequalities $a_n(i,m) < (1+i)^{\frac{1}{m}} \cdot 10^{\frac{n-1}{m}} \leq a_n(i,m) + 1$, one obtains that $a_n(i,m)^m < (1+i) \cdot 10^{n-1} \leq [a_n(i,m)+1]^m$. For $i \in \{1, \ldots, 8\}$ the middle term is the smallest number with n digits and first digit $1+i$, and for $i = 9$ it is the smallest number with $n+1$ digits. The result follows from the definition of the m-th power counting sequence. □

Remark 1. It is clear that the sequences $a_n(9,m) = \left\lceil 10^{\frac{n}{m}} \right\rceil - 1, \quad n \in N_+$, also describe the largest numbers whose m-th powers have n digits and are found in Sloane [25] under the sequences A049416 if $m = 2$ and A061439 if $m = 3$.

The integer sequences $\{a_n(i,m)\}$ are useful because they permit to count the number of m-th powers with at most n digits and first digit $i \in \{1,...,9\}$, a sequence denoted here $\{b_n(i,m)\}$. These integer sequences are new and not contained in Sloane [25].

Proposition 1. *The number of m-th powers with at most n digits and first digit $i \in \{1,...,9\}$ satisfies the following explicit formulas:*

$$\begin{aligned} b_n(1,m) &= 1 + \sum_{k=2}^{n} [a_k(1,m) - a_{k-1}(9,m)], \\ b_n(i,m) &= \sum_{k=1}^{n} [a_k(i,m) - a_k(i-1,m)], \quad i \in \{2,...,9\}. \end{aligned} \tag{2}$$

Proof. This follows immediately from the defining properties of the sequences $\{a_n(i,m)\}$. □

It is now possible to describe the *exact probability distribution* of the first digit of m-th integer powers with at most n digits, denoted and equal to

$$p_n(i,m) = \frac{b_n(i,m)}{a_n(9,m)}, \quad i \in \{1,...,9\}. \tag{3}$$

The formula (3) follows from (2) and the fact that

$$a_n(9,m) = \sum_{i=1}^{9} b_n(i,m) \tag{4}$$

is the largest number whose m-th power has n digits (see Remark 1 above).

Calculations with m-th power integer sequences $\{n^m\}$ show that the frequency of the first significant digit $d \in \{1,...,9\}$ decreases with increasing d, which suggests a strong relationship with Benford's law. Recall some relevant notions. For a sequence $S = \{a_n \,|n \in N_+\}$ consider the subsets $S_d = \{n \,|n \in S$ and the first digit of n is $d\}$. Then S is called a *Benford sequence* provided

$$\lim_{N\to\infty} \frac{\text{card}\,(S_d < N)}{\text{card}\,(S < N)} = \log\left(1 + \frac{1}{d}\right), \quad d \in \{1,...,9\}. \tag{5}$$

One knows that S is a Benford sequence if and only if the logarithmic sequence $\{\log a_n\}$ is uniformly distributed mod 1 (e.g. Diaconis [9]). A necessary and sufficient condition for this is the criterion of Weyl [28]:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h \log a_n} = 0 \;\text{ for all }\; h \in N_+. \tag{6}$$

Moreover, if S is a Benford sequence, then one has necessarily (Fejer's Theorem):

$$\lim_{n\to\infty} n \cdot \ln \left\{ \frac{a_{n+1}}{a_n} \right\} = \infty. \tag{7}$$

These results are presented in Kuipers and Niederreiter (1974).

Unfortunately, since $\lim_{n\to\infty} n\cdot\ln\{(n+1)^m/n^m\} = m\cdot\ln\{(1+1/n)^n\} = m < \infty$, the power sequences $\{n^m\}$ are not Benford sequences. However, as m increases, the exact probability distribution of the first significant digit approaches very closely Benford's law according to the following precise result.

Theorem 1. *The asymptotic distribution of the first digit of m-th power sequences as the number of digits goes to infinity is given by*

$$\lim_{n\to\infty} p_n(i,m) = P(i,m) := \frac{(1+i)^{\frac{1}{m}} - i^{\frac{1}{m}}}{10^{\frac{1}{m}} - 1}, \qquad i \in \{1,...,9\}, \quad m \geq 2. \tag{8}$$

Proof. As the number n of digits increases one has with great accuracy

$$b_n(i,m) \approx \left[(1+i)^{\frac{1}{m}} - i^{\frac{1}{m}}\right] \cdot \frac{10^{\frac{n}{m}} - 1}{10^{\frac{1}{m}} - 1}, \qquad i \in \{1,...,9\}, \quad a_n(9,m) \approx 10^{\frac{n}{m}} - 1, \tag{9}$$

which implies immediately (8). □

By increasing exponent m, the probability distribution (8) is closely approximated by Benford's law and is reached asymptotically because

$$\lim_{m\to\infty} P(i,m) = \log\left(1 + \frac{1}{i}\right), \quad i \in \{1,...,9\}. \tag{10}$$

$100p_{20}(i,2)$	$\lvert p_{20}(i,2) - P(i,2)\rvert$	$\lvert p_{20}(i,2) - \log(1+i^{-1})\rvert$	$100p_{30}(i,3)$	$\lvert p_{30}(i,3) - P(i,3)\rvert$	$\lvert p_{30}(i,3) - \log(1+i^{-1})\rvert$
19.156	$5.298\cdot 10^{-10}$	$1.095\cdot 10^{-1}$	22.515	$4.146\cdot 10^{-10}$	$7.588\cdot 10^{-2}$
14.699	$0.017\cdot 10^{-10}$	$0.291\cdot 10^{-1}$	15.794	$2.305\cdot 10^{-10}$	$1.815\cdot 10^{-2}$
12.392	$4.561\cdot 10^{-10}$	$0.010\cdot 10^{-1}$	12.573	$2.785\cdot 10^{-10}$	$0.080\cdot 10^{-2}$
10.918	$3.411\cdot 10^{-10}$	$0.123\cdot 10^{-1}$	10.618	$1.511\cdot 10^{-10}$	$0.927\cdot 10^{-2}$
9.870	$2.715\cdot 10^{-10}$	$0.195\cdot 10^{-1}$	9.281	$3.692\cdot 10^{-10}$	$1.363\cdot 10^{-2}$
9.077	$5.083\cdot 10^{-10}$	$0.238\cdot 10^{-1}$	8.299	$0.460\cdot 10^{-10}$	$1.605\cdot 10^{-2}$
8.448	$3.182\cdot 10^{-10}$	$0.265\cdot 10^{-1}$	7.542	$2.127\cdot 10^{-10}$	$1.743\cdot 10^{-2}$
7.935	$4.876\cdot 10^{-10}$	$0.282\cdot 10^{-1}$	6.937	$2.207\cdot 10^{-10}$	$1.822\cdot 10^{-2}$
7.505	$1.560\cdot 10^{-10}$	$0.293\cdot 10^{-1}$	6.440	$5.468\cdot 10^{-10}$	$1.865\cdot 10^{-2}$
$100p_{50}(i,5)$	$\lvert p_{50}(i,5) - P(i,5)\rvert$	$\lvert p_{50}(i,5) - \log(1+i^{-1})\rvert$	$100p_{100}(i,10)$	$\lvert p_{100}(i,10) - P(i,10)\rvert$	$\lvert p_{100}(i,10) - \log(1+i^{-1})\rvert$
25.423	$3.610\cdot 10^{-10}$	$4.680\cdot 10^{-2}$	27.720	$0.541\cdot 10^{-10}$	$2.383\cdot 10^{-2}$
16.590	$0.529\cdot 10^{-10}$	$1.019\cdot 10^{-2}$	17.128	$4.841\cdot 10^{-10}$	$0.481\cdot 10^{-2}$
12.614	$5.948\cdot 10^{-10}$	$0.120\cdot 10^{-2}$	12.581	$2.456\cdot 10^{-10}$	$0.087\cdot 10^{-2}$
10.296	$5.446\cdot 10^{-10}$	$0.605\cdot 10^{-2}$	10.011	$1.815\cdot 10^{-10}$	$0.320\cdot 10^{-2}$
8.760	$1.697\cdot 10^{-10}$	$0\text{-}842\cdot 10^{-2}$	8.347	$2.997\cdot 10^{-10}$	$0.429\cdot 10^{-2}$
7.660	$3.602\cdot 10^{-10}$	$0.966\cdot 10^{-2}$	7.177	$5.378\cdot 10^{-10}$	$0.482\cdot 10^{-2}$
6.829	$0.544\cdot 10^{-10}$	$1.030\cdot 10^{-2}$	6.307	$6.085\cdot 10^{-10}$	$0.508\cdot 10^{-2}$
6.177	$1.399\cdot 10^{-10}$	$1.062\cdot 10^{-2}$	5.633	$0.576\cdot 10^{-10}$	$0.518\cdot 10^{-2}$
5.650	$1.880\cdot 10^{-10}$	$1.074\cdot 10^{-2}$	5.096	$0.461\cdot 10^{-10}$	$0.520\cdot 10^{-2}$
$100\cdot p_{1000}(i,100)$	$\lvert p_{1000}(i,100) - P(i,100)\rvert$	$\lvert p_{1000}(i,100) - \log(1+i^{-1})\rvert$	$100\cdot p_{10000}(i,1000)$	$\lvert p_{10000}(i,1000) - P(i,1000)\rvert$	$\lvert p_{10000}(i,1000) - \log(1+i^{-1})\rvert$
29.861	$3.022\cdot 10^{-10}$	$2.419\cdot 10^{-3}$	30.079	$0.807\cdot 10^{-9}$	$2.422\cdot 10^{-4}$
17.564	$6.929\cdot 10^{-10}$	$0.453\cdot 10^{-3}$	17.605	$5.667\cdot 10^{-9}$	$0.450\cdot 10^{-4}$
12.505	$3.543\cdot 10^{-10}$	$0.111\cdot 10^{-3}$	12.495	$1.755\cdot 10^{-9}$	$0.114\cdot 10^{-4}$
9.724	$5.950\cdot 10^{-10}$	$0.334\cdot 10^{-3}$	9.694	$2.255\cdot 10^{-9}$	$0.336\cdot 10^{-4}$
7.962	$3.165\cdot 10^{-10}$	$0.434\cdot 10^{-3}$	7.922	$2.241\cdot 10^{-9}$	$0.435\cdot 10^{-4}$
6.743	$9.079\cdot 10^{-10}$	$0.481\cdot 10^{-3}$	6.699	$3.820\cdot 10^{-9}$	$0.480\cdot 10^{-4}$
5.849	$8.237\cdot 10^{-10}$	$0.500\cdot 10^{-3}$	5.804	$3.408\cdot 10^{-9}$	$0.500\cdot 10^{-4}$
5.166	$4.827\cdot 10^{-10}$	$0.506\cdot 10^{-3}$	5.120	$3.102\cdot 10^{-9}$	$0.505\cdot 10^{-4}$
4.626	$4.075\cdot 10^{-10}$	$0.504\cdot 10^{-3}$	4.581	$2.123\cdot 10^{-9}$	$0.503\cdot 10^{-4}$

Table 1: Distribution of first digit of integer powers and Benford's law

Table 1 illustrates numerically the obtained results. The exact probability distribution (3) is calculated for $n = 10m$, for various exponents m, and compared with the asymptotic distributions (8) and (10).

References

[1] P.C. Allaart, An invariant-sum characterization of Benford's law, *Journal of Applied Probability*, **34** (1997), 288-291.

[2] P. Becker, Patterns in listings of failure-rate and MTTF values and listings of other data, *IEEE Transactions on Reliability*, **R-31** (1982), 132-134.

[3] A. Berger, L.A. Bunimovich, T.P. Hill, One-dimensional dynamical systems and Benford's law, In: *Transactions of the American Mathematical Society* (2002).

[4] F. Benford, The law of anomalous numbers, *Proceedings of the American Philosophical Society*, **78** (1938), 551-572.

[5] J. Brown, R. Duncan, Modulo one uniform distribution of the sequence of logarithms of certain recursive sequences, *Fibonacci Quarterly*, **8** (1970), 482-486.

[6] B. Buck, A. Merchant, S. Perez, An illustration of Benford's first digit law using alpha decay half lives, *European Journal of Physics*, **14** (1993), 59-63.

[7] J. Burke, E. Kincanon, Benford's law and physical constants: the distribution of initial digits, *American Journal of Physics*, **59** (1991), 952.

[8] D. Cohen, T. Katz, Prime numbers and the first digit phenomenon, *Journal of Number Theory*, **18** (1984), 261-268.

[9] P. Diaconis, The distribution of leading digits and uniform distribution mod 1, *Annals of Probability*, **5** (1977), 72-81.

[10] G.A. Gottwald, On the nature of Benford's law, *Physica A: Statistical Mechanics and its Applications*, **303**, No. 3-4 (2002), 387-396.

[11] T.P. Hill, Base-invariance implies Benford's law, *Proceedings of the American Mathematical Society*, **123** (1995), 887-895.

[12] T.P. Hill, The significant-digit phenomenon, *Amer. Math. Monthly*, **102** (1995), 322-326.

[13] T.P. Hill, A statistical derivation of the significant-digit law, *Statistical Science*, **10** (1995), 354-363.

[14] T.P. Hill, Benford's law, *Encyclopedia of Mathematics Supplement*, Volume **1** (1997), 102.

[15] T.P. Hill, The first digit phenomenon, *The American Scientist*, **10** (1998), 354-363.

[16] E. Janvresse, T. de la Rue, From uniform distribution to Benford's law, In: *Publication de l'UMR 6085*, Université de Rouen (2003).

[17] L. Kuipers, H. Niederreiter, *Uniform Distribution of Sequences*, J. Wiley (1973).

[18] E. Ley, On the peculiar distribution of the U.S. Stock Indices Digits, *The American Statistician*, **50** (1996), 311-313.

[19] S. Newcomb, Note on the frequency of use of the different digits in natural numbers, *American Journal of Mathematics*, **4** (1881), 39-40.

[20] M. Nigrini, A taxpayer compliance application of Benford's law, *Journal of the American Taxation Association*, **1** (1996), 72-91.

[21] M. Nigrini, I've got your number, *Journal of Accountancy*, **187** (1999), 79-83.

[22] M. Nigrini, L. Mittermaier, The use of Benford's law as an aid in analytical procedures, *A Journal of Practice and Theory*, **16** (1997), 52-67.

[23] R.S. Pinkham, On the distribution of first significant digits, *Annals of Mathematical Statistics*, **32** (1961), 1223-1230.

[24] P. Schatte, On Benford's law for continued fractions, *Math. Nachr.*, **148** (1990), 137-144.

[25] N.J.A. Sloane, Ed. *The On-Line Encyclopedia of Integer Sequences*, Published electronically at:
www.research.att.com/∼njas/sequences (2003).

[26] C. Tolle, J. Budzien, R. LaViolette, Do dynamical systems follow Benford's law?, *Chaos*, **10** (2000), 331-337.

[27] W. Vogt, Benford's Gesetz: Steuer- und Budgetsündern auf der Spur-Zahlen lügen nicht, *Schweizer Versicherung*, **9** (2000), 27-29.

[28] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, *Mathematische Annalen*, **77** (1916), 313-352.

[29] R.E. Whitney, Initial digits for the sequence of primes, *American Mathematical Monthly*, **79** (1972), 150-152.