Benford's Law to Base g of Order r in the Sense of a Certain Density

Kenji Nagasaka*
Hosei University

Benford's law is originally an empirical probabilistic law of first significant digits from a set of numerous real numbers, first reported in literature by Simon Newcomb [8]. This empirical law was rediscovered by Frank Benford [1], and since then called Benford's law after his name.

Frank Benford collects many kinds of data and considers the distribution of first significant digits. Then his conjecture is as follows: "The first significant digit from numerical tables does not occur with equal frequency, but the earlier digits appears more often than the latter's."

This conjecture may also be stated as a theoretical model which predicts the proportion of entries beginning with first digit k, for $k = 1, 2, \cdots, 9$ is well approximated by $\log_{10}(1 + 1/k)$. In this note, we restrict ourselves to infinite sequences of positive integers.

Let $U = (u_n)_{n=1}^{\infty}$ be a sequence of positive integers and $A_N(U; k)$ be the counting function of U for $k = 1, 2, \ldots, 9$ defined by

$$A_N(U; k) = \#\{i \le N; \text{the first digit of } u_i \text{ is equal to } k\},$$

where $\#S$ denotes the number of elements of the set S .

$U = (u_n)_{n=1}^{\infty}$ is said to obey Benford's law to base 10 if, for every $k = 1, 2, \ldots, 9$

$$\lim_{N\to\infty} \frac{A_N(U; k)}{N} = \log_{10}(k + 1) - \log_{10} k = \log_{10}\left(\frac{k + 1}{k}\right).$$

It is known that the sequence $(3^n)_{n=1}^{\infty}$ obeys Benford's law to base 10. But if we consider this sequence $(3^n)_{n=1}^{\infty}$ to base 3, then all first digits of $(3^n)_{n=1}^{\infty}$ to base 3 are equal to one. For example, 3 to base 10 is represented by $(10)_3$ to base 3 and 27 by $(1000)_3$ to base 3. Thus we are led to consider Benford's law to base g, where g in an integer greater than 2, since the distribution of first digits is in question.

Let Δ_r be a block of length r defined by

$$\Delta_r = (k_1, k_2, \ldots, k_r),$$

for $k_1 = 1, 2, \ldots, 9$ and $k_i = 0, 1, \ldots, 9$ for $i = 2, 3, \ldots, r$ and $A_N(U; \Delta_r)$ be the counting function of U for which the first r consecutive digits of u_i coincide with the block Δ_r for $1 \le i \le N$.

Definition. $U = (u_n)_{n=1}^{\infty}$ is said to obey Benford's law to base 10 **of order** r if, for every block Δ_r of length r

$$\lim_{N\to\infty} \frac{A_N(U; \Delta_r)}{N} = \log_{10}(k_1k_2\cdots k_r + 1) - \log_{10}(k_1k_2\cdots k_r),$$

*nagasaka@hosei.ac.jp

where $k_1k_2\cdots k_r = k_1 \cdot 10^{r-1} + k_2 \cdot 10^{r-2} + \cdots + k_r$ and further at least for one block Δ_{r+1} of length $r+1$,

$$\lim_{N\to\infty} \frac{A_N(U;\Delta_{r+1})}{N} \neq \log_{10}(k_1k_2\cdots k_{r+1}+1) - \log_{10}(k_1k_2\ldots k_{r+1}),$$

where $\Delta_{r+1} = (k_1, k_2, \ldots, k_{r+1})$, for $k_1 = 1, 2, \ldots, 9$ and $k_i = 0, 1, \ldots, 9$ for $i = 2, 3, \ldots, r+1$.

Theorem 1. $U = (u_n)_{n=1}^{\infty}$ obeys Benford's law to base 10 of order r if and only if $\log_{10} U = (\log_{10} u_n)_{n=1}^{\infty}$ satisfies

$$\lim_{N\to\infty} \frac{A_N(\log_{10} U; I(\Delta_r))}{N} = \mu_0(I(\Delta_r)) , \tag{1}$$

for every Δ_r of length r and for at least one block Δ_{r+1} of length $r+1$,

$$\lim_{N\to\infty} \frac{A_N(\log_{10} U; I(\Delta_{r+1}))}{N} \neq \mu_0(I(\Delta_{r+1})) ,$$

where μ_0 denotes the Lebesgue measure and $I(\Delta_r)$ is the interval of the form

$$[\ \{\log_{10}(k_1k_2\cdots k_r)\},\ \{\log_{10}(k_1k_2\cdots k_r+1)\}\ ) , \tag{2}$$

further $A_N(\log_{10} U; I(\Delta_r))$ is defined by

$$A_N(\log_{10} U; I(\Delta_r)) = \#\{\ i; 1 \le i \le N,\ \{\log_{10} u_i\} \in I(\Delta_r)\},$$

and for a real number x , $\{x\}$ denotes the fractional part of x , i.e. $\{x\} = x - \lfloor x \rfloor$ and $0 \le \{x\} < 1$.

A sequence of positive integers $U = (u_n)_{n=1}^{\infty}$ obeys strong Benford's law to base 10 if U obeys Benford's law of order r for every $r = 1, 2, \cdots$ The following Theorem is due to P. Diaconis [3] as his Theorem 1.

Theorem 2. $U = (u_n)_{n=1}^{\infty}$ obeys strong Benford's law to base 10 if and only if the sequence $\log_{10} U = (\log_{10} u_n)_{n=1}^{\infty}$ is uniformly distributed mod 1.

We can define analogously for $U = (u_n)_{n=1}^{\infty}$ to obey Benford' law of order r **to base** g.

Definition. $U = (u_n)_{n=1}^{\infty}$ is said to obey **strong Benford's law to base** g if U obeys Benford's law to base g of order r for every $r \ge 3$.

Definition. $U = (u_n)_{n=1}^{\infty}$ is said to obey **absolutely strong Benford's law** if U obeys strong Benford's law to every base $g \ge 3$.

Theorem 3. There exist sequences $U = (u_n)_{n=1}^{\infty}$ which obey strong Benford's law to base s and do not obey to base t if and only if there exist no positive integers m and n such that $s^m = t^n$.

Now we remark the existence of a sequence which obeys absolutely strong Benford's law, such as the sequence of Fibonacci numbers, which can be generalized easily by using Theorem 4.1 in [6] . For linear recurrence integer sequences of order an arbitrary $d \ge 2$, we can obtain almost the same results as above in Nagasaka [6], [7] and Shigeru Kanemitsu et. al [5].

Theorem 2 due to Diaconis can be generalized as:

Theorem 4. $U = (u_n)_{n=1}^{\infty}$ obeys strong Benford's law to base g in the sense of uniform density if and only if the sequence $\log_g U = (\log_g u_n)_{n=1}^{\infty}$ is well distributed mod 1.

For other densities, such as Dirichlet density, H_∞-density, supernatural density, the sequence of all positive integers obeys absolutely strong Benford's law in the sense of the above densities. [4], [2].

References

[1] F. Benford:*The law of anomalous numbers* , **Proc. Amer. Phil. Soc., 78** (1938), 551-572.

[2] D. I. A. Cohen:*An explanation of the first digit phenomenon* , **J. Combinatorial Theory, Ser. A 20** (1976), 367-370.

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

[4] B. J. Flehinger:*On the probability that a random integer has initial digit* A, **Amer. Math. Monthly, 73** (1966), 1056-1061.

[5] S. Kanemitse, K. Nagasaka, G. Rauzy and J.-S. Shiue: *On Benford's law: The first digit problem* , Probability Theory and Mathematical Statistics (Kyoto, 1986) , **Lecture Notes in Math., 1299**, Springer-Verlag, Berlin, New York (1988), 158-169.

[6] K. Nagasaka:*On Benford's Law* , **Annals of Inst. Statist. Math., 36-2, A** (1984), 337-352.

[7] K. Nagasaka, Shigeru Kanemitsu and Jau-Shyong Shiue:*Benford's law: The logarithmic law of first digit* , **Collq. Mah. Soc. J́anos Bolyai. 51**, Number Theory, Budapest, Hungary (1987), 361-391.

[8] S. Newcomb:*Note on the frequency of use of the different digits in natural numbers* , **Amer. J. Math., 4** (1881), 39-40.

[9] R. A. Raimi:*The first digits problem*, **Amer. Math. Monthly, 83** (1970), 521-538.