A NEW APPROACH TO THE FIRST DIGIT PHENOMENON

BY

DENNIS P. ALLEN, JR.

Abstract:- In this paper, we first show by extending the proof of B.J. Flehinger [1] *that the integers have the first digit property that the primes represented to the base ten also have the first digit property. We note that R. E. Whitney*[2] *has also proven this using the logarithmic matrix method of summability. We then abstract from these two proofs in view of the Peano Axioms to obtain a (new) definition of what it means to sum a sequence in the spirit of the Peano axioms for the positive integers (which includes Flehinger's and Whitney's methods as special cases) and conjecture any method of summation of this very general type assigns the limit* $\log_{10}((A+1)/A)$ *to the two sequences* s_n *and* t_n *where*

(a) $s_n = 1$ *if n has first digit equal to A else 0, and*

(b) $t_n = 1$ *if the* n^{th} *prime has first digit A else 0.*

The conjecture, if true, yields a new explanation of the first digit phenomenon by reducing it to its first cause: the basic well ordering of the positive integers.

Is there a meaningful way to answer the heuristic question, "What proportion of the prime numbers have initial digit less than or equal to A?" or "What is the probability that a prime number chosen at random has initial digit less than or equal to A?" This question was asked by B. J. Flehinger[1, 3] only with "prime number" replaced by "positive integer". Her question was occasioned by her interest in explaining a well known phenomenon, first reported by Frank Benford[4], that the proportion of numbers in tables of physical constants with first significant digit less than or equal to A is $\log_{10}(A+1)$. She reasoned that "the smallest population which contains the set of significant figures of all positive physical constants, past, present and future, must be the set of positive integers" and so that "the explanation for the logarithmic law should, therefore, lie in the properties of the set of integers as represented in a radix number system".

Flehinger went on to show that while, if n is a positive integer and $P_n^1(A)$ is the proportion of the positive integers less than or equal to n which has initial digit less than or equal to A, the $\lim_{n\to\infty} P_n^1(A)$ does not exist, that this sequence is Cesaro summable and we have

$$\lim_{k\to\infty} \liminf_{n\to\infty} P_n^k(A) = \lim_{k\to\infty} \limsup_{n\to\infty} P_n^k(A) = \log_{10}(A+1), \text{ where}$$

The Toth-Maatian Review, 3101 20th Street, Lubbock, TX 79410.
Volume 14, #3, 1999, pp. 6839-47.

$$P_n^k(A) = \frac{1}{n}\left(\sum_{m=1}^{n} P_m^{k-1}(A)\right) \text{ for } k = 2,3,4,\ldots$$

(Note that this limiting process is very dependent on the ≤ order of the integers which is so central to their definition in terms of the Peano Axioms [5].*) Her discussion is limited to the base ten number system, but the generalization to an arbitrary radix is straightforward. We will show that the primes "inherit" this property of the positive integers, that is, the answer to the question posed at the beginning is the proportion of the prime numbers having initial digit less than or equal to A is $\log_{10}(A+1)$, by modifying Flehinger's proof and making use of the prime number theorem with remainder [6]. More specifically, we will first show that if $\hat{P}_n^1(A)$ is the proportion of prime numbers less than or equal to N having initial digit less than or equal to A, that (here, following Flehinger, we change notation and write $Q^1(\alpha,A)$ for

$\lim_{j\to\infty} \hat{P}^1_{\alpha\cdot 10^j}(A)$ for clarity)

$$Q^1(\alpha,A) = \lim_{j\to\infty} \hat{P}^1_{\alpha\cdot 10^j}(A) = \begin{cases} 1 - \frac{9-A}{9\alpha} & 1 \le \alpha < A+1 \\ \frac{10A}{9\alpha} & A+1 \le \alpha < 10, \end{cases}$$

that is that while $\hat{P}_N^1(A)$ does not approach a limit as $N \to \infty$ either, that $\hat{P}^1_{\alpha\cdot 10^j}(A)$ does converge for every α with $1 \le \alpha < 10$. To accomplish this we apply the prime number thorem with remainder.

As is customary, we let $\pi(x)$ equal the number of primes less than or equal to x for x a positive real number. The prime number theorem with remainder states that

$$\pi(x) = x/\ln(x) + 0(x/\ln^2(x)),$$

---

* Of course, the order is implicit in the concept of measurement.

and so (considering only the main term x/ln(x) and reserving discussion of the error term until later) for $1 \le \alpha < A+1$, we have

$$Q^1(a,A) = \lim_{j \to \infty} \hat{P}^1_{\alpha \cdot 10^j}(A)$$

$$= \lim_{j\to\infty} \left\{ \frac{1}{\pi(a\cdot 10^j)} \left(\sum_{0\le k \langle j} \pi((A+1)\cdot 10^k) - \pi(10^k) \right.\right.$$

$$\left.\left. + \left(\pi(a\cdot 10^j) - \pi(10^j)\right) \right) \right\}$$

$$= \lim_{j\to\infty} \left\{ \frac{\ln(\alpha) + j\ln(10)}{\alpha\cdot 10^j} \left(\sum_{0\le k \langle j} \left(\frac{(A+1)10^k}{\ln(A+1) + k\ln(10)} - \frac{10^k}{k\ln(10)} \right) \right.\right.$$

$$\left.\left. + \left(\frac{\alpha\cdot 10^j}{\ln(\alpha) + j\ln(10)} - \frac{10^j}{j\ln(10)} \right) \right) \right\}$$

$$= \lim_{j\to\infty} \left\{ \frac{j\ln(10)}{\alpha\cdot 10^j} \left(\sum_{1\le k \langle j} \left(\frac{(A+1)10^k}{k\ln(10)} - \frac{10^k}{k\ln(10)} \right) + \right.\right.$$

$$\left.\left. + \left(\frac{\alpha\cdot 10^j}{j\ln(10)} - \frac{10^j}{j\ln(10)} \right) \right) \right\}$$

$$= \lim_{j\to\infty} \left\{ \frac{1}{\alpha} \left(A\left(\sum_{1\le k \langle j} \frac{j}{k\cdot 10^{j-k}} \right) + (\alpha - 1) \right) \right\} = 1 - \frac{9-A}{9\alpha}$$

The predicted value, since

$$\lim_{j \to \infty} \sum_{k=1}^{j-1} \frac{j}{k \cdot 10^{j-k}} = \frac{1}{9}.$$

We see that we will have shown that $Q^1(\alpha,A)$ has the predicted value if we can successfully argue that the error term $0(x/\ln^2(x))$ can be neglected in the computation of this quantity. But we have

$$\frac{x/\ln(x)}{x/\ln(x) + 0(x/\ln^2(x))} = \frac{1}{1 + 0(1/\ln(x))} \to 1 \text{ as } x \to \infty \cdot$$

So the error term will not affect our computation.

In the case where $A+1 \leq \alpha < 10$, we have (ignoring the error term again)

$$Q^1(\alpha, A) = \lim_{j \to \infty} \hat{P}^1_{\alpha \cdot 10^j}(A) = \frac{10A}{\alpha}$$

in the same way as we computed $Q^1(\alpha,A)$ in the case $1 \leq \alpha < A + 1$. So $Q^1(\alpha,A)$ has the predicted value for $1 \leq \alpha < 10$ as we wished to prove.

Now, following Flehinger, we observe that, for all k,

$\lim_{j\to\infty} \hat{P}^k_{\alpha \cdot 10^j}(A)$ exists where $\hat{P}^k_n(A) = \frac{1}{n} \sum_{m=1}^{n} \hat{P}^{k-1}_m(A)$ for all $k \rangle 1$ and we set this limit equal to

$Q^k(\alpha, A)$. Clearly, we have

$$\liminf_{n\to\infty} P^k_n(A) = \min_{1 < \alpha < 10} Q^k(\alpha, A) \text{, and}$$

$$\limsup_{n\to\infty} P^k_n(A) = \max_{1 < \alpha < 10} Q^k(\alpha, A) .$$

Therefore, we can complete the proof by showing that $\lim_{k\to\infty} Q^k(\alpha,A) = \log_{10}(A+1)$ for $1 \leq \alpha < 10$.

But it is easy to see that we can derive the recursive formula

$$Q^k(\alpha,A) = \frac{1}{\alpha}\left\{ \frac{1}{9}\int_1^{10} Q^{k-1}(\beta,A)\,\delta\beta \;+\; \int_1^{\alpha} Q^{k-1}(\beta,A)\,\delta\beta \right\}$$

in the same way as it is derived in Flehinger's paper. From this point on the proof is the same, and so we conclude that

$$\lim_{k\to\infty} Q^k(\alpha, A) = \log_{10}(A+1)$$

Which completes the proof.

Note that it follows that the probability that a prime will have initial digit A is just $\log_{10}((A+1)/A)$ for $1 \leq A \leq 9$.

Philosophical Discussion. The logarithmic law is surprising, precisely because, when the positive integers are represented to the base 10, one tends to think of them merely as finite strings of digits, that is, as finite strings from the alphabet {0,1,...,9}. But this is to confuse two completely different semigroups: the semigroup of finite strings from {0,1,...,9} whose first element is not zero (under concatenation) and the semigroup of positive integers. The former is a right ideal of the free semigroup of all strings on this alphabet [7] while the latter is given by the Peano Axioms [5] which stress the unique well order of the positive integers. Of course, the primes are defined recursively (in that multiplication is defined recursively), that is, in terms of the natural order of these numbers. And not just defined recursively, but deeply recursively in the sense that the primes are fundamental in the structure of the positive integers which is why the fundamental theorem of arithmetic has that name. Thus, it is, perhaps, not too surprising that the primes also satisfy the logarithm law.

We are indebted to Professor R. A. Raimi for the information that in 1972 R. E. Whitney published a paper [2] in which he uses the logarithmic matrix method of summability to define a density function and then he uses this density function to show that the relative logarithmic density of the primes with initial digit A in the prime is $\log_{10}((A+1)/A)$. This logarithmic matrix method of summability is stronger than the Cesaro method used here([8], p. 524).

Page 6844.

The author believes that he sees the reason why both Whitney's and Flehinger's methods of summability give the result $\log_{10}((A+1)/A)$. Both regular methods of summation are strong enough to force convergence and each sums the sequence in question in the spirit of the Peano axioms. Here, we say that a regular method of summation sums a sequence in the spirit of the Peano axioms if when it is applied to a sequence $\{a_n\}$, it gives another sequence $\{b_n\}$ which is related to $\{a_n\}$ as follows:

(a) the sequence $\{b_n\}$ should converge or at least be closer to convergence than before and in this case, repeated application should force convergence if only in the limit,

(b) for each natural number j, b_j should depend only on $a_1, a_2 \ldots, a_j$ just as j is obtained from Peano Axioms by applying the successor function (j -1) times to 1,

(c) in computing b_2, a_2 should be weighted no more heavily* than a_1 (both weights having a positive sign) just as 2 is the successor of 1 and is (with respect to the axioms) no more important than 1,

(c) in computing b_3, a_3 should be weighted no more heavily than a_2 which in turn should be weighted no more heavily than a_1 just as 3 is obtained from 2 and 2 from 1 by applying the successor function and hence (with respect to the axioms) 3 is no more important than 2, and 2 no more than 1; etc.

It may be helpful to translate this definition into the usual language of summability in terms of infinite matrices. First, we are only talking about regular methods of summation, that is, those which when applied to a convergent sequence yield the usual limit.

Condition (a) of the definition seems to be clear, and we would only add that it is necessary to allow "convergence if only in the limit" because Flehinger's method of summation is repeated Cesaro summation.

---

*Although perhaps less heavily. For many purposes, the integers become less important as they become larger.

Condition (b) of our definition is also clear, but conditions (c) and (d) may not be. We would explain them as follows: for each positive integer n there exists weights $W_{n,1}$, $W_{n,2}\ldots,W_{n,n}$ such that

$$b_n = W_{n,1}a_1 + W_{n,2}a_2 + \ldots + W_{n,n}a_n \text{ and with } W_{n,1}\geq W_{n,2}\geq\ldots\geq W_{n,n}>0.$$

Thus, if we set $W_{i,j} = 0$ for $j > i > 0$, we have that $W_{i,j}$ is a lower triangular Toeplitz matrix.

Clearly, whenever the limit of a sequence is obtained by a regular method, this limit cannot depend on any finite number of the values of the terms of the sequence. Thus we can weaken this condition for the Toeplitz matrix to read: there exists $k > 0$ so that we have $W_{n,m}\geq W_{n,m+1}\geq W_{n,m+2}\geq \ldots W_{n,n} > 0$ whenever $n \geq m > k$, and still have equivalence.

The author conjectures that if either of the two sequences $a_n^{(1)}$ or $a_n^{(2)}$ is summed in the spirit of the Peano axioms, then the result is $\log_{10}((A+1)/A)$, where $a_n^{(1)} = 1$ if the n^{th} integer has initial digit A when represented to the base 10 and otherwise equals 0, and where $a_n^{(2)} = 1$ if the n^{th} prime has initial digit A when represented to the base 10 and otherwise equals 0.

The truth of the first part of this conjecture means that if positive integers are used for counting with the 1's before the 2's, the 2's before the 3's, … , and the 8's before the 9's (this condition, of course, rules out, for example, the case of seven digit telephone numbers which never begin with a 1) or for addition in such a way that the set of all integers used is closed under addition provided that the sum is not too large (where all sufficiently large integers are unimportant) or if a set of decimal numbers is used which is closed under addition of numbers whose sums are not too large and with the numbers having a bounded number of non-zero digits, then the subset of important integers has (approximately) the first digit property of Benford. (Here in the case of decimal numbers with a bounded number of non-zero digits, the subset of important integers has as members those which are obtained from at least one important decimal number by multiplying it by 10^k where k is the smallest nonnegative integer making this product also an integer.) However, if the numbers are not used this way (in the spirit of the Peano axioms), the subset of important numbers need not satisfy Benford's law. As an example of this, the set of all finite strings of digits from the alphabet 0,1, …, 9 whose initial digit is not 0 is (as mentioned above) a semigroup under concatenation in which the natural probability of a string having initial digit A is 1/9, and it can be embedded in a type of semidirect product of the semigroup of positive integers with itself (which also is a semigroup) in such a way that the multiplication is preserved [9]. Thus the higher algebraic operations need not preserve Benford's law at all; they can lead far from the original Peano axioms defining their constituents.

Such a set of integers or decimal numbers closed under addition provided the sum is important is called a Benford set. It is typically generated by the numbers used in the course of calculating a table of numbers. Here, by a table of numbers, we mean a finite sequence of numbers selected from some finite parent population where the same number can occur more than once in the table. An example would be the length of the rivers in a certain country in miles. It often happens, as in this case, that little or no information about the values of the table can be inferred except that the finite population out of which the sequence values are to be selected may have certain properties (e.g. it may be the important numbers of a Benford set). In this situation, the probability that a given member of the parent population will be assigned to a term in the sequence may be 1/n where n is the cardinality of the finite parent population. Then the set of sequence values will tend to inherit the statistical properties of the parent population if both are sufficiently large. In particular, if the parent population is the important members of a Benford set, as is the case of the example mentioned, then the table will (approximately) satisfy Benford's law. Thus we avoid the hypothesis of scale invariance which, however, comes in the back door in the example since it is the fact that the rivers could also be measured in yards or feet or so on, for example, that allows us to conclude that probability a table entry is a given member of the parent population is 1/n.

We note that the Peter Schatte [10] has proven the first result of this paper using the theory of the uniform distribution of sequences. His techniques do not seem to help in the proving of either part of our conjecture which, if true, seems to explain the first digit phenomenon pretty well.

As a final comment, we note that Bumby and Ellentuck [11] have given an explanation of the logarithm law using finitely additive measures. Following Pinkham [3,8], they use a modified notion of scale invariance to obtain this law. Although very interesting, their methods do not seem to lend themselves very easily to showing that the primes satisfy this law too; the summability approach appears to be the more powerful of the two.

REFERENCES

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

[2] R. E. Whitney, Initial Digits for the Sequence of Primes, Amer. Math Monthly **79** (1972), pp. 150-152.

[3] D. E. Knuth, **The Art of Computer Programming, Vol. 2, Seminumerical Algorithms**, Addison-Wesley, London, 1969, pp. 219-228.

[4] Frank Benford, The law of anomalous numbers, Proc. Amer. Philos. Soc., **78** (1938), 551-572.

[5] R. L. Wilder, **Introduction to the Foundations of Mathematics**, second ed., John Wiley & Sons, New York, 1967.

[6] H. Davenport, **Multiplicative Number Theory**, Markham, Chicago, 1967.

[7] A. H. Clifford and G. Preston, **The Algebraic Theory of Semigroups**, Vol. 1 (1961), Vol. 2 (1967), Amer. Math. Soc., Math Surveys 7.

[8] R. A. Raimi, The First Digit Problem, Amer. Math. Monthly **83** (1976), pp. 521-538.

[9] Dennis Allen, Jr., On a Characterization of the Nonregular Set of Primes, J. Computer System Sciences, **2** (1968) 464-467.

[10] Peter Schatte, On H_∞ - Summability and the Uniform Distribution of Sequences, Math. Nachr. **113** (1983) 237-243.

[11] R. Bumby and E. Ellentuck, Finitely additive measures and the first digit problem, Fund. Math., **65** (1969), 33-42.

Dennis P. Allen, Jr.
P.O. Box 41
Spring Lake, MI 49456-0041
e-mail: allend@internationalaid.org
First submitted: January 8, 1999

THE TOTH-MAATIAN PRESS, 1998

I.S.S.N. 0740-7864

Printed in the U.S.A.