Randomness in Number Theory: Leading Digits Distribution

IGL Project Report, Spring 2016*

Zhaodong Cai, Matthew Faust, Shunping Xie
Junxian Li (Team Leader),
A.J. Hildebrand (Faculty Mentor)†

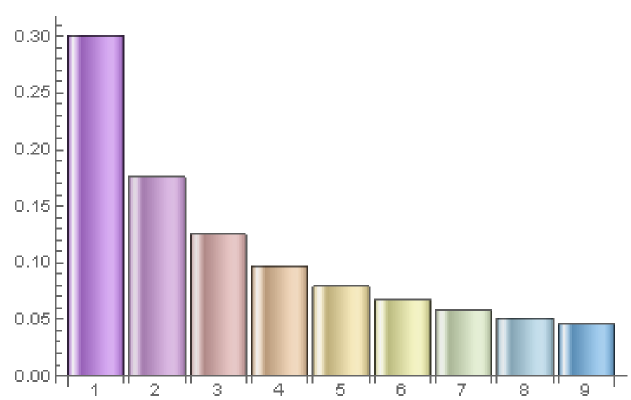
Illinois Geometry Lab
University of Illinois at Urbana-Champaign

May 10, 2016

If you pick a Fibonacci number at random, what are the chances that it begins with a 1 (or a 2, or a 9)? Surprisingly, the answer is not 1/9 as one might expect. In fact, around 30.1% of Fibonacci numbers begin with a 1, about 17.6% begin with a 2, and only around 4.6% begin with a 9. More generally, the probability that a Fibonacci number begins with digit d is given by

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right), \quad d = 1, 2, \ldots, 9. \tag{1}$$

This peculiar first-digit distribution, depicted below, is known as Benford's Law [1], [3].



*An expanded version of this report is being prepared for possible publication.

†Email: ajh@illinois.edu

Benford's Law has been found to apply to many real world data sets such as populations of world cities, masses of planets, stock price indices, and numbers in accounting statements, and it has important applications to fraud detection.

From a mathematical point of view, Benford's Law is closely connected with the theory of *uniform distribution modulo* 1 [4]. Diaconis [2] used this connection to prove rigorously that Benford's Law holds (in an appropriate asymptotic sense) for certain classes of mathematical sequences, for example, the powers of 2, the factorials, and the Fibonacci numbers. That is, in each of these sequences, the digit d occurs as leading digit with the asymptotic frequency given by (1).

The general goal of this project, which started in Fall 2015, is to investigate the leading digit distribution of mathematical sequences such as the Fibonacci numbers more closely, both theoretically (e.g., by exploiting connections with the theory of uniform distribution modulo 1) and experimentally (e.g., through large scale computations or visualizations with Mathematica). In particular, we are interested in the degree of global and local "randomness" in the sequences of leading digits.

Last semester, we focused on getting familiar with the underlying theory, and with relevant topics from the theory of uniform distribution modulo 1 and the theory of continued fractions. We also carried out numerical computations that revealed surprising patterns in the behavior of leading digit sequences and suggested some tantalizing conjectures.

This semester we continued working on both the theoretical and the experimental sides of the problem. We were able to prove our main conjecture from Fall 2015, the "Local Benfordness Conjecture," and some related results. On the experimental side we used the Campus Computing Cluster to significantly extend our numerical computations and develop evidence for further conjectures.

Preliminary results of this work have been presented at the *Joint Mathematics Meetings*, Seattle, WA, Jan. 8, 2016 (Zhaodong Cai, Yuda Wang); the *UI Undergraduate Research Symposium*, April 21, 2016 (Zhaodong Cai, Matthew Faust, Shunping Xie); and the *Rose-Hulman Undergraduate Mathematics Conference*, Terre Haute, IN, April 22-23, 2016 (Zhaodong Cai, Matthew Faust, Shunping Xie). A paper is being prepared for possible publication..

References

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

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

[3] T.P. Hill, *The significant-digit phenomenon*, American Math. Monthly 102 (1995), 322–327.

[4] L. Kuipers and L. Niederreiter, *Uniform distribution of sequences*, Dover Publications, 2006.