Local and Global Randomness in the Leading Digits of Arithmetic Sequences

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

ILLINOIS UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN   Illinois Geometry Lab

University of Illinois: Undergraduate Research Symposium, Urbana, IL, April 21, 2016

Benford's Law

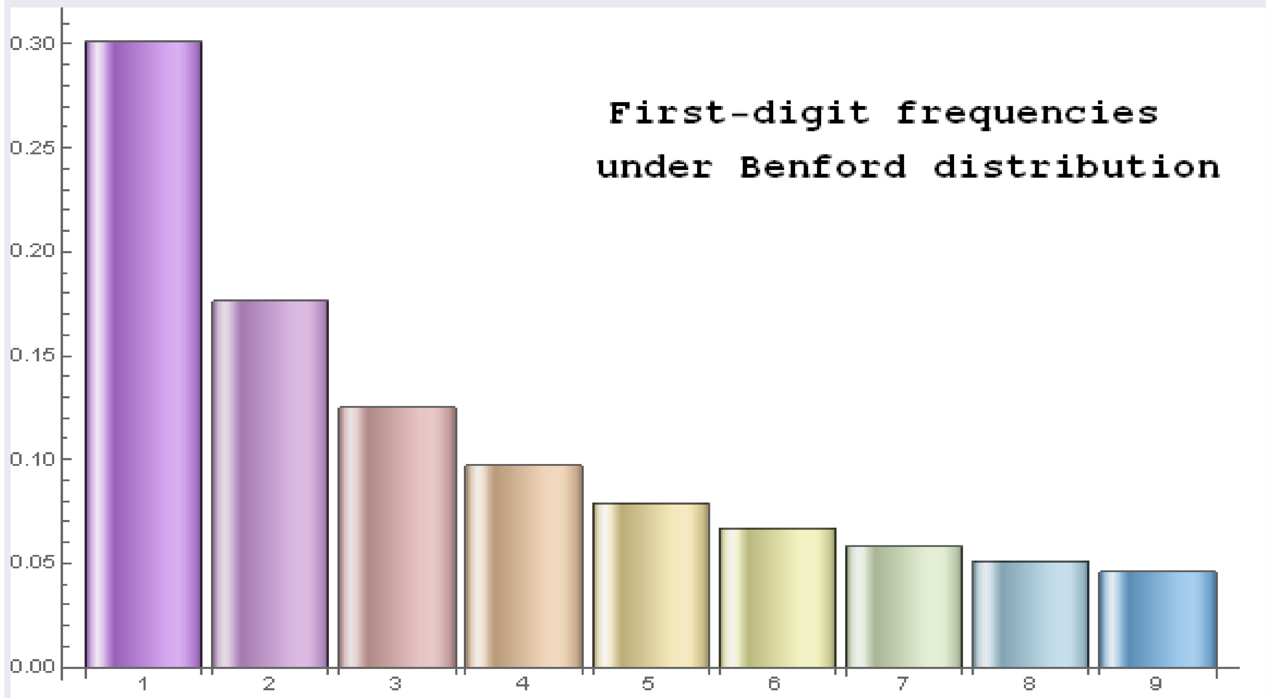
*A sequence $\{x_n\}$ is said to be **Benford distributed** if the probability of having first digit d is*

$$P(\text{first digit } d) = \log_{10}(1 + \frac{1}{d})$$

Real world examples satisfying (approx.) Benford's Law:

- Populations of US cities
- Areas of countries
- Physical constants
- File sizes in Linux file system
- Numbers in US tax returns

Frank Benford (1883–1948)



Benford's Law in Mathematics: Fibonacci Numbers

Actual counts and Benford predictions ($10^9 \log_{10} \frac{d+1}{d}$) for the leading digits of the **first billion Fibonacci numbers**.

First Digit	Actual count	Benford Prediction	Difference
1	301029995	301029995.66	-0.66
2	176091265	176091259.06	$+5.94$
3	124938730	124938736.61	-6.61
4	96910014	96910013.01	$+0.99$
5	79181254	79181246.05	$+7.95$
6	66946785	66946789.63	-4.63
7	57991942	57991946.98	-4.98
8	51152529	51152522.45	$+6.55$
9	45757486	45757490.56	-4.56

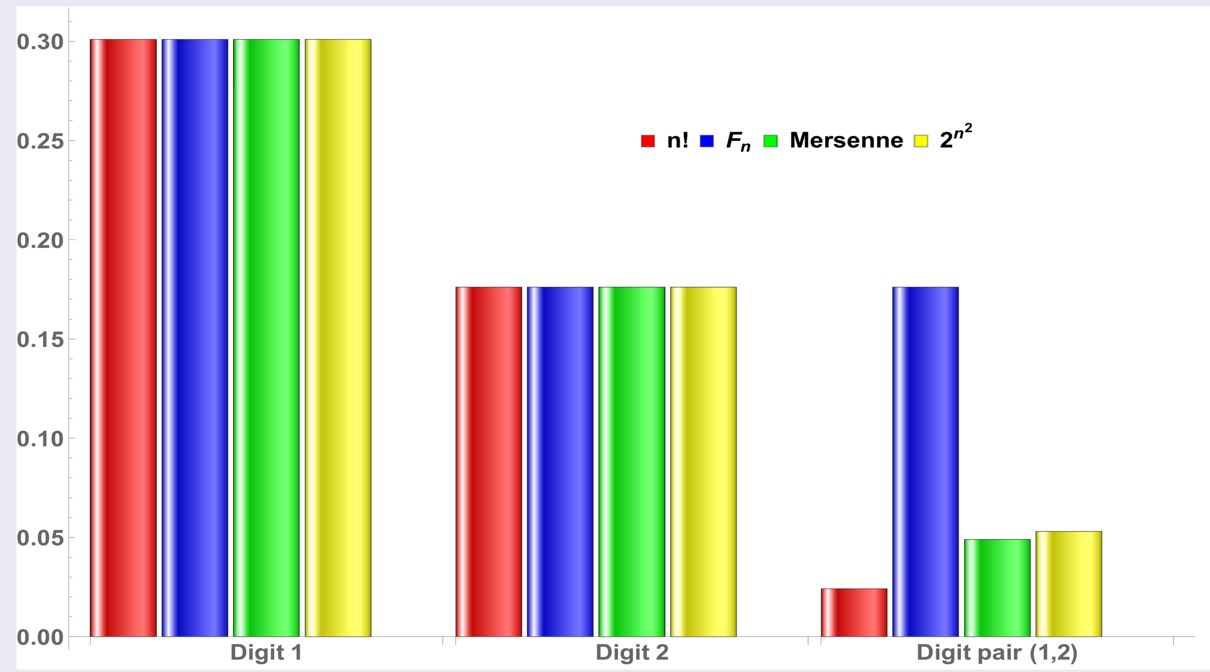
Leading Digits of Fibonacci Numbers

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, 10946, 17711, 28657, 46368, 75025, 121393, 196418, 317811, 514229, 832040, 1346269, 2178309, 3524578, 5702887, 9227465, 14930352, 24157817, 39088169, 63245986, 102334155, 165580141, 267914296, 433494437, 701408733, 1134903170, 1836311903, 2971215073, 4807526976, 7778742049, 12586269025, 20365011074, 32951280099, 53316291173, 86267571272, 139583862445, 225851433717,

Problem: How random are such first digit sequences?

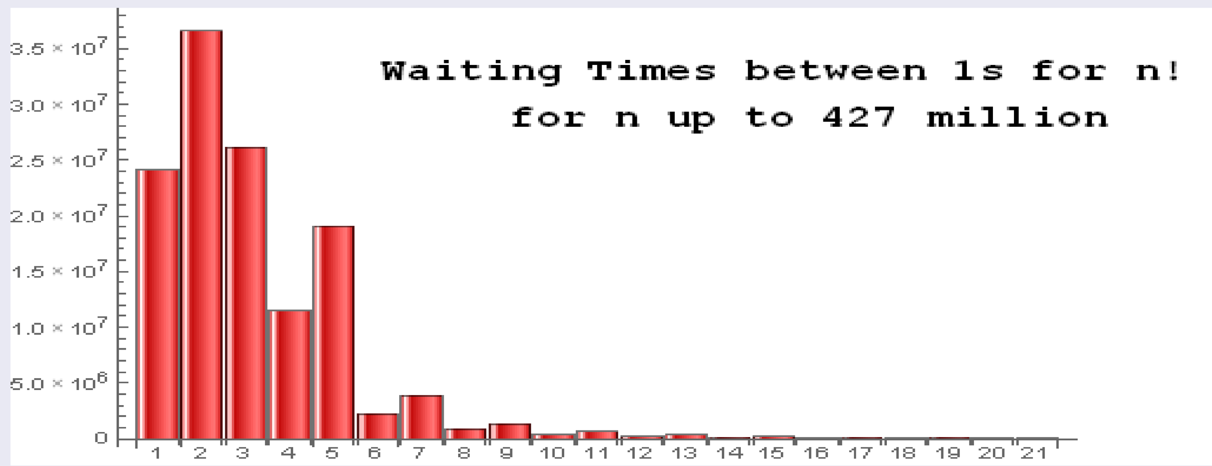
Local Benfordness: Distributions of Pairs

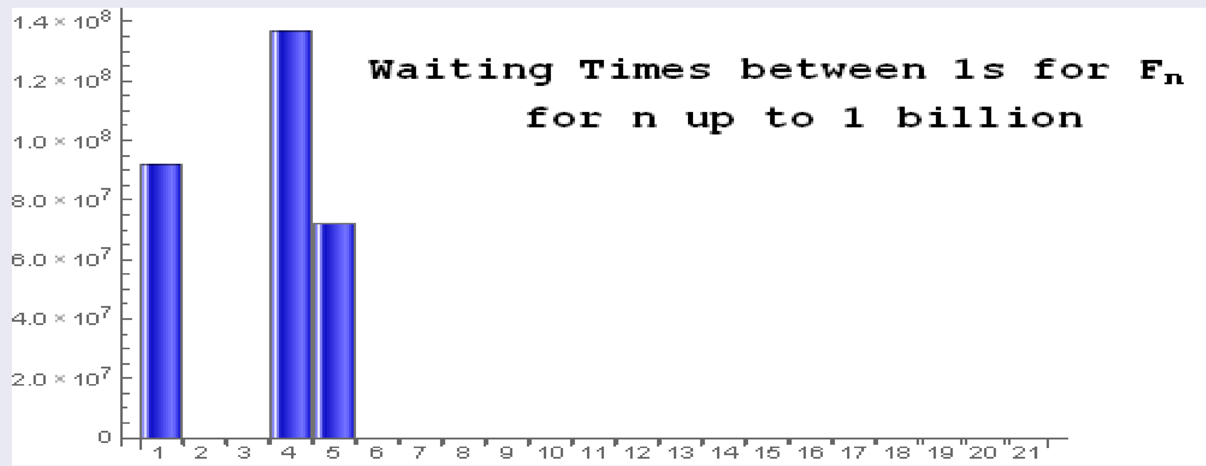
Sequence	Freq of 1	Freq of 2	Freq of (1,2)
$n!$	0.301058	0.176136	0.024134
2^n	0.301030	0.176091	0.176091
F_n	0.301030	0.176091	0.176091
Mersenne	0.300997	0.176092	0.048968
2^{n^2}	0.301053	0.176085	0.052987
Benford	0.301030	0.176091	0.053009

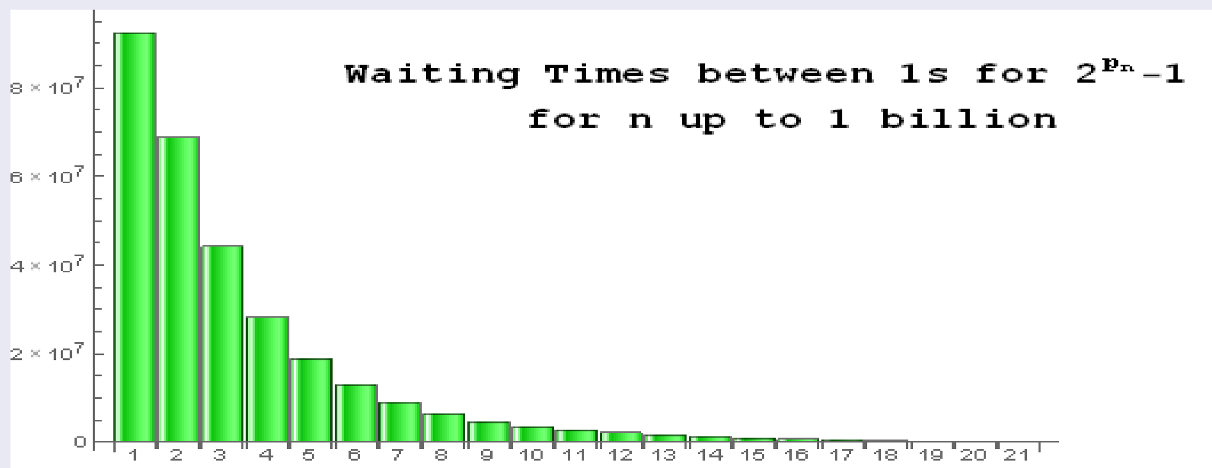


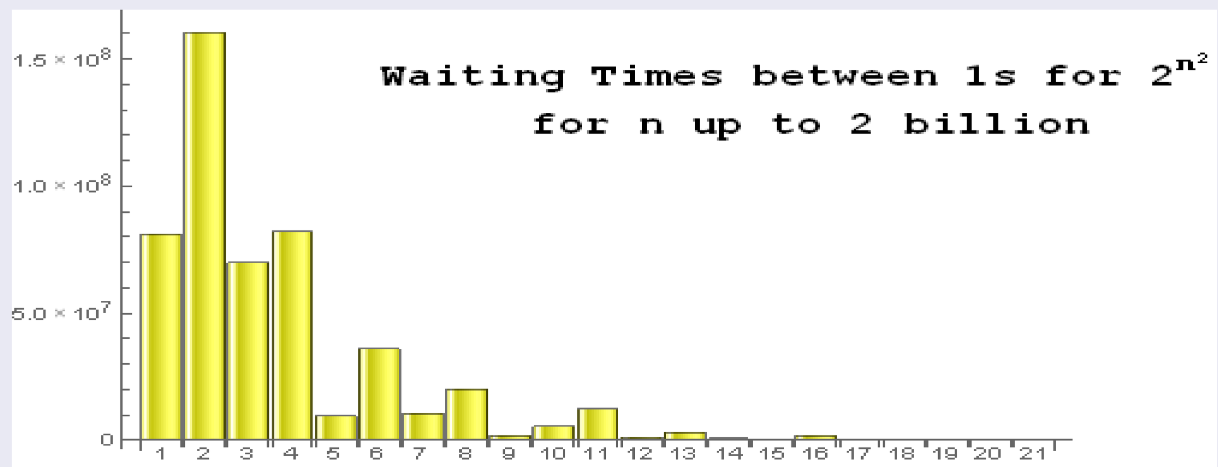
Numerical data based on 100,000,000 terms suggest that the leading digits of $n!$, 2^n, F_n, and Mersenne numbers are Benford distributed, but pairs of digits are not independent. By contrast, pairs of leading digits of 2^{n^2} are independent and Benford distributed.

Local Benfordness: Waiting Times between digit 1's









Theoretical Results

Definition (Locally Benford)

A sequence is called locally Benford of order d if the leading digits of $(a_{n+1}, \ldots, a_{n+k})$ have the same asymptotic distribution as k independent Benford distributions, for all $k \leq d$, but not for $k > d$.

Theorem (Local Benfordness Theorem)

Let $a_n = a^{n^d}(1 + o(1))$, where $\log_{10} a \notin \mathbb{Q}$ and d is a positive integer. Then a_n is locally Benford of order d.

Example:

- F_n, 2^n and Mersenne numbers are locally Benford of order 1, i.e., the distributions of leading digits for these sequences satisfy Benford's Law, but the distributions of pairs of leading digits are not independent.
- 2^{n^2} is locally Benford of order 2, i.e., the distribution of leading digits satisfies Benford's Law and pairs of leading digits are independent, but triples of leading digits are not independent.

Waiting Times

Theorem (First-digit Waiting Times)

(i) *If $a_n = a^n(1 + o(1))$ where $\log_{10} a \notin \mathbb{Q}$, the sequence has bounded waiting times between digits.*

(ii) *If $\frac{a_{n+1}}{a_n} = n^k(1 + o(1))$ for some $k > 0$, the sequence has unbounded waiting times between digits.*

Example:

- 2^n, F_n have bounded waiting time between 1's
- $n!$ has unbounded waiting times between 1's

References

- F. Benford, The Law of Anomalous Numbers, Proc. Amer. Philos. Soc., 78(4) (1938), pp. 551- 572.
- P. Diaconis, The Distribution of Leading Digits and Uniform Distribution Mod 1, Annals of Probability, Vol. 5, No. 1 (1977), pp. 72-81.
- T.P. Hill, The Significant-digit Phenomenon, Amer. Math. Monthly, Vol 102, No.4 (1995), pp. 322-327.
- L. Kuipers, L. Niederreiter, Uniform Distribution of Sequences, Dover Publications (2006).