The amusing and excellent law of Benford

Principles of Complex Systems
Course CSYS/MATH 300, Fall, 2009

Prof. Peter Dodds

Dept. of Mathematics & Statistics
Center for Complex Systems :: Vermont Advanced Computing Center
University of Vermont

Outline

The law of first digits

Benford's Law:

- First observed by Simon Newcomb[2] in 1881
  "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered by Frank Benford in 1938.
- Newcomb almost always noted but Benford gets the stamp
- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$

  for numbers in base b

The law of first digits

Benford's Law:

- First observed by Simon Newcomb [2] in 1881
  "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered by Frank Benford in 1938.
- Newcomb almost always noted but Benford gets the stamp
- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
  for numbers in base b

The law of first digits

Benford's Law:

- First observed by Simon Newcomb [2] in 1881
  "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered by Frank Benford in 1938.
- Newcomb almost always noted but Benford gets the stamp
-

$$P(\text{first digit} = d) \propto \log_b(d + 1/d)$$

for numbers in base b

The law of first digits

Benford's Law:

- First observed by Simon Newcomb [2] in 1881
  "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered by Frank Benford in 1938.
- Newcomb almost always noted but Benford gets the stamp
-
$$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
  for numbers in base b

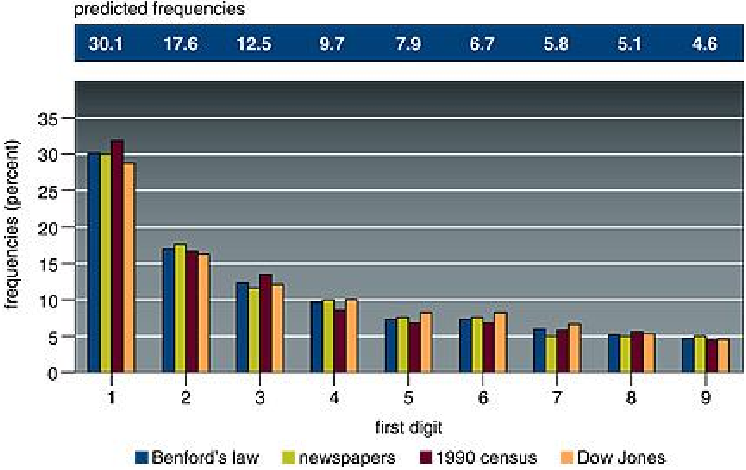
Benford's Law—The law of first digits

Observed for

- Fundamental constants (electron mass, charge, etc.)
- Utilities bills
- Numbers on tax returns
- Death rates
- Street addresses
- Numbers in newspapers

Benford's Law

Real data



From 'The First-Digit Phenomenon' by T. P. Hill (1998) [1]

Essential story

- $$P(\text{first digit} = d) \propto \log_b(d + 1/d)$$
- $$P(\text{first digit} = d) \propto \log_b\left(\frac{d+1}{d}\right)$$
- $$P(\text{first digit} = d) \propto \log_b(d+1) - \log_b(d)$$
- So numbers are distributed uniformly in log-space:
$$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$
- Independent of actual base and units of measurement.
- Power law distributions at work again... ($\gamma = 1$)

Essential story

- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
- $$P(\text{first digit} = d) \propto \log_b \left(\frac{d+1}{d}\right)$$
- $$P(\text{first digit} = d) \propto \log_b (d + 1) - \log_b (d)$$
- So numbers are distributed uniformly in log-space:
$$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$
- Independent of actual base and units of measurement.
- Power law distributions at work again... ($\gamma = 1$)

Essential story

- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
- $$P(\text{first digit} = d) \propto \log_b \left(\frac{d+1}{d}\right)$$
- $$P(\text{first digit} = d) \propto \log_b (d + 1) - \log_b (d)$$
- So numbers are distributed uniformly in log-space:
$$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$
- Independent of actual base and units of measurement.
- Power law distributions at work again... ($\gamma = 1$)

Essential story

- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
- $$P(\text{first digit} = d) \propto \log_b \left(\frac{d+1}{d}\right)$$
- $$P(\text{first digit} = d) \propto \log_b (d + 1) - \log_b (d)$$
- So numbers are distributed uniformly in log-space:

$$P(\ln x)\, \mathrm{d}(\ln x) \; \propto 1 \cdot \mathrm{d}(\ln x) \; = x^{-1}\, \mathrm{d}x$$

- Independent of actual base and units of measurement.
- Power law distributions at work again... ($\gamma = 1$)

Essential story

- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
- $$P(\text{first digit} = d) \propto \log_b \left(\frac{d+1}{d}\right)$$
- $$P(\text{first digit} = d) \propto \log_b (d+1) - \log_b (d)$$
- So numbers are distributed uniformly in log-space:
$$P(\ln x)\,\mathrm{d}(\ln x) \;\propto 1 \cdot \mathrm{d}(\ln x) \;= x^{-1}\,\mathrm{d}x$$
- Independent of actual base and units of measurement.
- Power law distributions at work again... ($\gamma = 1$)

Essential story

- $$P(\text{first digit} = d) \propto \log_b (d + 1/d)$$
- $$P(\text{first digit} = d) \propto \log_b \left(\frac{d+1}{d}\right)$$
- $$P(\text{first digit} = d) \propto \log_b (d + 1) - \log_b (d)$$
- So numbers are distributed uniformly in log-space:

$$P(\ln x)\,\mathrm{d}(\ln x) \;\propto 1 \cdot \mathrm{d}(\ln x) \;= x^{-1}\,\mathrm{d}x$$

- Independent of actual base and units of measurement.
- Power law distributions at work again... ($\gamma = 1$)

A different Benford

Not to be confused with Benford's Law of controversy:

A different Benford

Not to be confused with Benford's Law of controversy:

- "Passion is inversely proportional to the amount of real information available."

A different Benford

Not to be confused with Benford's Law of controversy:

- "Passion is inversely proportional to
  the amount of real information available."

Gregory Benford, Sci-Fi writer & Astrophysicist

References I

T. P. Hill.
The first-digit phenomenon.
American Scientist, 86:358–, 1998.

S. Newcomb.
Note on the frequency of use of the different digits in natural numbers.
American Journal of Mathematics, 4:39–40, 1881.
pdf (⊞)