The Amusing Law of Benford

Principles of Complex Systems
CSYS/MATH 300, Spring, 2013 | #SpringPoCS2013

Prof. Peter Dodds
@peterdodds

Department of Mathematics & Statistics | Center for Complex Systems | Vermont Advanced Computing Center | University of Vermont

Outline

The law of first digits

Benford's Law: (⊞)

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  for certain sets of 'naturally' occurring numbers in base b
- Around 30.1% of first digits are '1', compared to only 4.6% for '9'.
- First observed by Simon Newcomb [2] in 1881 "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered in 1938 by Frank Benford (⊞).
- Newcomb almost always noted but Benford gets the stamp.

The law of first digits

Benford's Law: (⊞)

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$

  for certain sets of 'naturally' occurring numbers in base b
- Around 30.1% of first digits are '1', compared to only 4.6% for '9'.
- First observed by Simon Newcomb [2] in 1881 "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered in 1938 by Frank Benford (⊞).
- Newcomb almost always noted but Benford gets the stamp.

The law of first digits

Benford's Law: (⊞)

- $$P(\text{first digit} = d) \propto \log_b\left(1 + \frac{1}{d}\right)$$
  for certain sets of 'naturally' occurring numbers in base b
- Around 30.1% of first digits are '1', compared to only 4.6% for '9'.
- First observed by Simon Newcomb [2] in 1881 "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered in 1938 by Frank Benford (⊞).
- Newcomb almost always noted but Benford gets the stamp.

The law of first digits

Benford's Law: (⊞)

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$

  for certain sets of 'naturally' occurring numbers in base b
- Around 30.1% of first digits are '1', compared to only 4.6% for '9'.
- First observed by Simon Newcomb [2] in 1881 "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered in 1938 by Frank Benford (⊞).
- Newcomb almost always noted but Benford gets the stamp.

The law of first digits

Benford's Law: (⊞)

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$

  for certain sets of 'naturally' occurring numbers in base b
- Around 30.1% of first digits are '1', compared to only 4.6% for '9'.
- First observed by Simon Newcomb [2] in 1881 "Note on the Frequency of Use of the Different Digits in Natural Numbers"
- Independently discovered in 1938 by Frank Benford (⊞).
- Newcomb almost always noted but Benford gets the stamp.

Benford's Law—The Law of First Digits

Observed for

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

- Cited as evidence of fraud (⊞) in the 2009 Iranian elections.

Benford's Law—The Law of First Digits

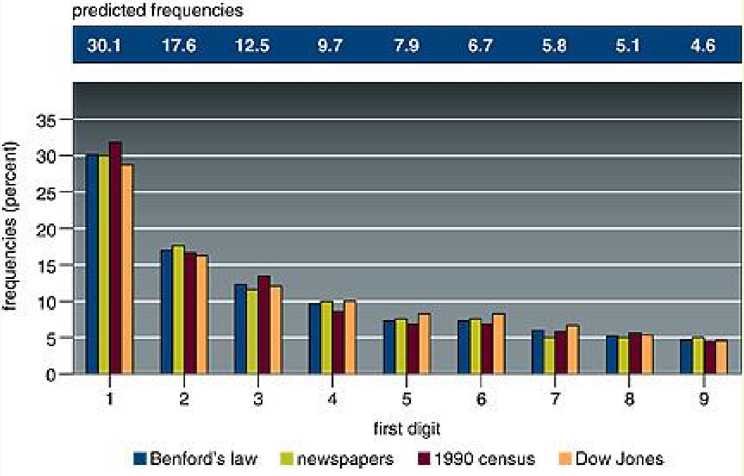
Observed for

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

- Cited as evidence of fraud (⊞) in the 2009 Iranian elections.

Benford's Law

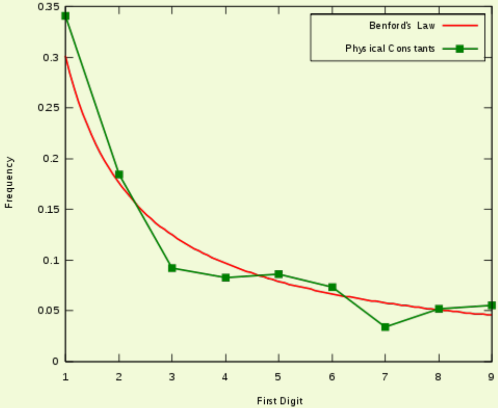
Real data:



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

Benford's Law

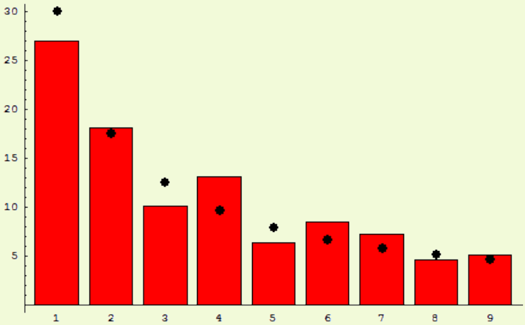
Physical constants of the universe:



Taken from here (⊞).

Benford's Law

Population of countries:



Taken from here (⊞).

Essential story

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$

- Observe this distribution if numbers are distributed uniformly in log-space:

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

- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Essential story

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  $$\propto \log_b \left(\frac{d+1}{d}\right)$$

- Observe this distribution if numbers are distributed uniformly in log-space:
  $$P(\ln x)\, \mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\, \mathrm{d}x$$

- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Essential story

- $$P(\text{first digit} = d) \propto \log_b\left(1 + \frac{1}{d}\right)$$
  $$\propto \log_b\left(\frac{d+1}{d}\right)$$
  $$\propto \log_b(d+1) - \log_b(d)$$
- Observe this distribution if numbers are distributed uniformly in log-space:
  $$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$
- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Essential story

- $$P(\text{first digit} = d) \propto \log_b\left(1 + \frac{1}{d}\right)$$
$$\propto \log_b\left(\frac{d+1}{d}\right)$$
$$\propto \log_b(d+1) - \log_b(d)$$

- Observe this distribution if numbers are distributed uniformly in log-space:

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

- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Essential story

- $$P(\text{first digit} = d) \propto \log_b\left(1 + \frac{1}{d}\right)$$
  $$\propto \log_b\left(\frac{d+1}{d}\right)$$
  $$\propto \log_b(d+1) - \log_b(d)$$

- Observe this distribution if numbers are distributed uniformly in log-space:
  $$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$

- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Essential story

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  $$\propto \log_b \left(\frac{d+1}{d}\right)$$
  $$\propto \log_b (d+1) - \log_b (d)$$
- Observe this distribution if numbers are distributed uniformly in log-space:
  $$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$
- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

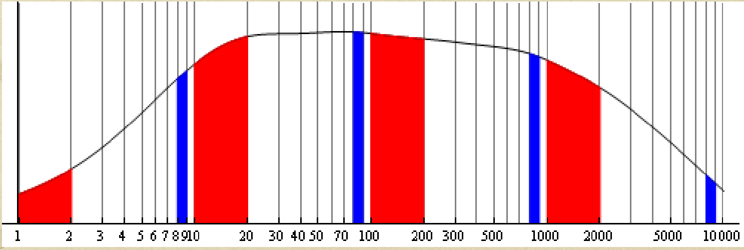
Essential story

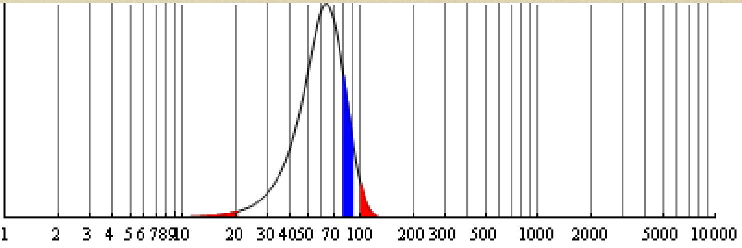
- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  $$\propto \log_b \left(\frac{d+1}{d}\right)$$
  $$\propto \log_b (d+1) - \log_b (d)$$

- Observe this distribution if numbers are distributed uniformly in log-space:
  $$P(\ln x)\, \mathrm{d}(\ln x) \;\propto 1 \cdot \mathrm{d}(\ln x) \;= x^{-1}\, \mathrm{d}x$$

- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Benford's law





Taken from here (⊞).

References I

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

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