The Amusing Law of Benford

Principles of Complex Systems | @pocsvox
CSYS/MATH 300, Fall, 2015 | #FallPoCS2015

Prof. Peter Dodds | @peterdodds

Dept. of Mathematics & Statistics | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

These slides are brought to you by:

Outline

Benford's Law—The Law of First Digits

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  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, according to Stigler's Law of Eponymy.

Benford's Law—The Law of First Digits

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

  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, according to Stigler's Law of Eponymy.

Benford's Law —The Law of First Digits

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  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, according to Stigler's Law of Eponymy.

Benford's Law—The Law of First Digits

- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  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, according to Stigler's Law of Eponymy.

Benford's Law—The Law of First Digits

- $$P(\text{first digit} = d) \propto \log_b\left(1 + \frac{1}{d}\right)$$
  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, according to Stigler's Law of Eponymy.

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

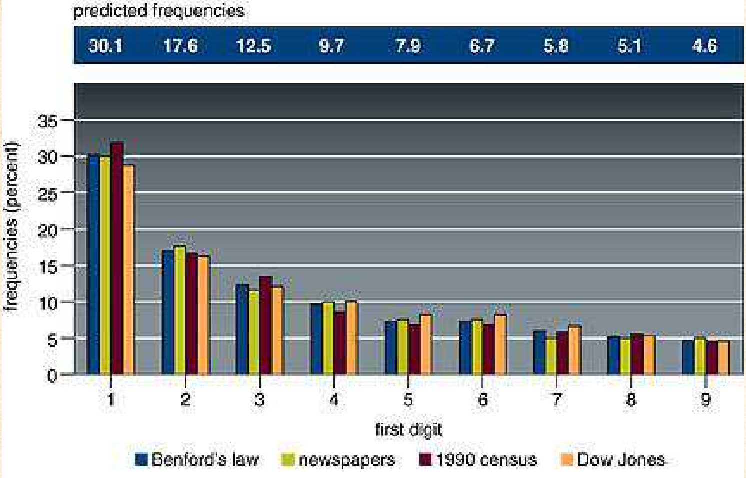
Observed for

- 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—The Law of First Digits

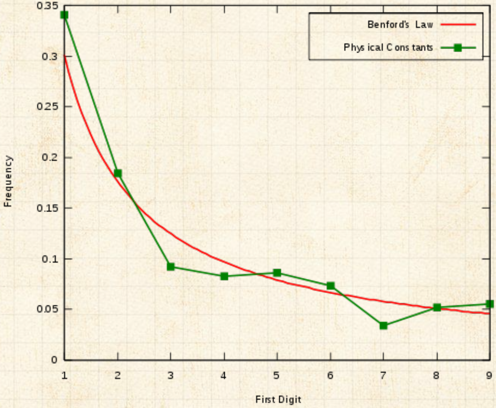
Real data:



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

Benford's Law—The Law of First Digits

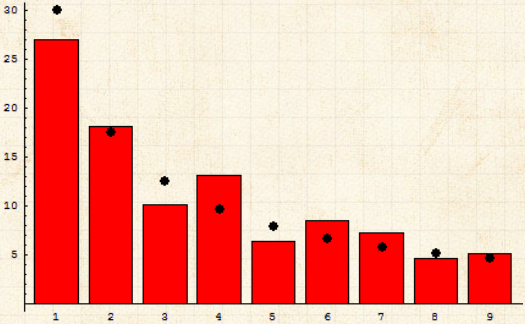
Physical constants of the universe:



Taken from here.

Benford's Law—The Law of First Digits

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:

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

- 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:

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

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

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

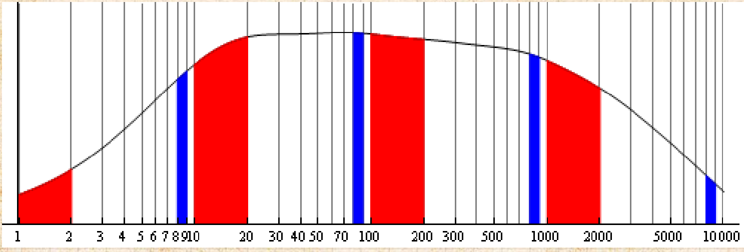
- Observe this distribution if numbers are distributed uniformly in log-space:
$$P(\ln x)\,\mathrm{d}(\ln x) \propto 1 \cdot \mathrm{d}(\ln x) = x^{-1}\,\mathrm{d}x$$

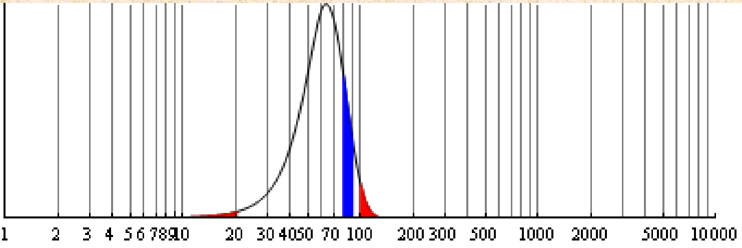
- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

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.

On counting and logarithms:

- Earlier: Listen to Radiolab's "Numbers.".
- Now: Benford's Law.

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