The Amusing Law of Benford

Principles of Complex Systems | @pocsvox
CSYS/MATH 300, Fall, 2016 | #FallPoCS2016

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


THE SUN
THE MANIFESTO
THE GOLDEN AGE OF REDUCTIONISM
THE WELL OF SCALES
THE LAWS OF DECAYING POWER
THE UNEXPECTED TRANSFIGURATION
MINISTRY OF RANDOM WALKS
THE RICH GET RICHER
THE DREAM OF EFFICIENT LANGUAGE


1 2
5 4 3
6 7 8
9
THE LAW OF
DIGITS FIRST


THE LAW OF
DIGITS FIRST

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 [3] 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 [3] 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 [3] 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 [3] 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 [3] 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

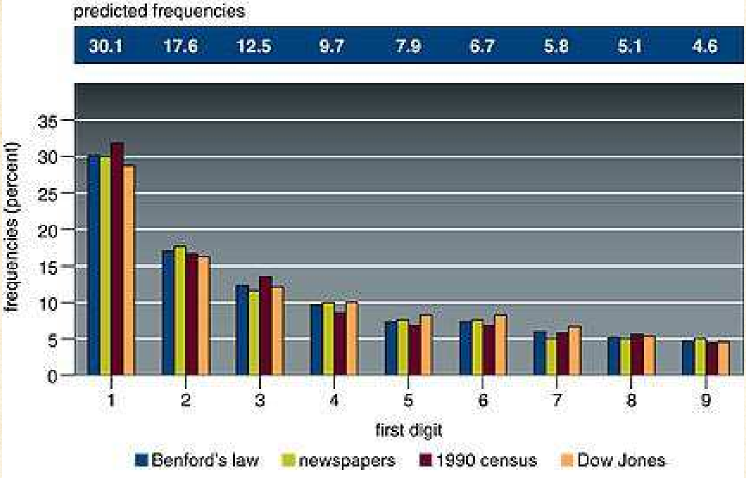
Observed for

- 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

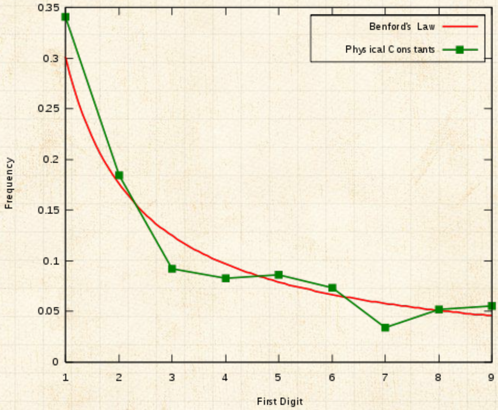
Real data:



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

Benford's Law—The Law of First Digits

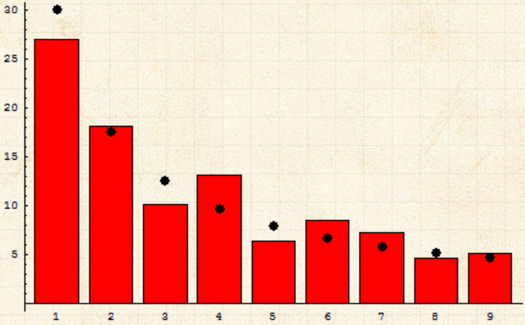
Physical constants of the universe:



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:

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

- 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:

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

- 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:

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

- 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:

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

- 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:
  $$P(\square\square\, x)\,\mathrm{d}(\square\square\, x) \propto 1 \cdot \mathrm{d}(\square\square\, x) = x^{-1}\,\mathrm{d}x$$
- Power law distributions at work again...
- Extreme case of $\gamma \simeq 1$.

Essential story

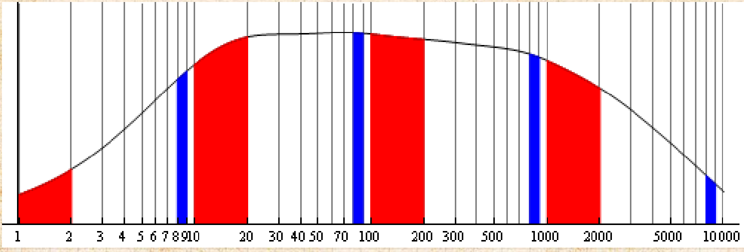
- $$P(\text{first digit} = d) \propto \log_b \left(1 + \frac{1}{d}\right)$$
  $$\propto \log_b \left(\frac{d+1}{d}\right)$$
  $$\propto \log_b (d+1) - \log_b (d)$$

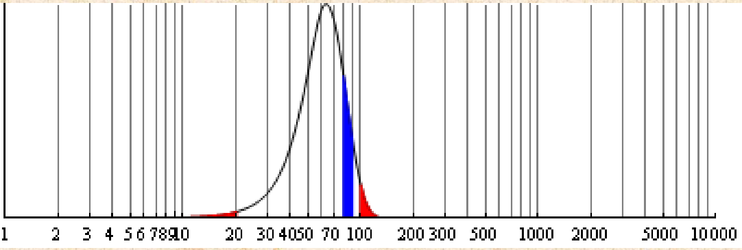
- Observe this distribution if numbers are distributed uniformly in log-space:

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

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

Benford's law





Taken from here.

"Citations to articles citing Benford's law: A Benford analysis"
Tariq Ahmad Mir,
Preprint available at
http://arxiv.org/abs/1602.01205,
2016. [2]

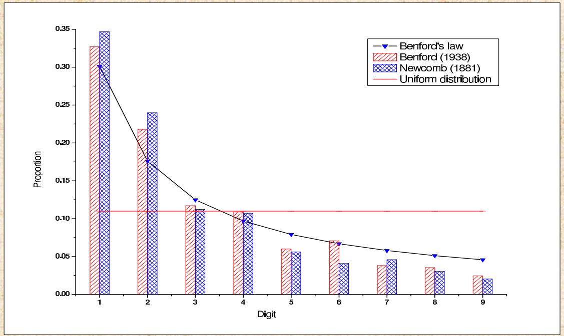


Fig. 1: The observed proportions of first digits of citations received by the articles citing FB and SN on September 30, 2012. For comparison the proportions expected from BL and uniform distributions are also shown.

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] T. A. Mir.
Citations to articles citing Benford's law: A Benford analysis, 2016.
Preprint available at
http://arxiv.org/abs/1602.01205. pdf

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