BenFord's Law

Anurag Dutta
1anuragdutta@gmail.com

2021-11-15

ACKNOWLEDGEMENT

I am overwhelmed in all humbleness and gratefulness to acknowledge at my depth to all those who have helped me to put these ideas, well above the level of simplicity and into something concrete.

I would like to express my special thanks of gratitude to my pedagogue as well as my University for giving me the golden opportunity to do this wonderful project on the topic - "*Reconnaissance of Election Data*" , which also helped me in doing a lot of Research and I came to know about so many new things. I am really thankful to them.

Any attempt at any level can't be satisfactorily completed without the support and guidance of my parents and friends on whom I am really thankful.

Contents

1 History

The discovery of this law had it's initiation in the late 1800 when Simon Newcomb, an astronomer noticed that in logarithm tables, the earlier pages (that started with 1) were much more worse than the other pages. Newcomb proposed a law that the probability of a single number N being the first digit of a number was equal to $log(N+1)-log(N)$. Later in the mid 1900s, Frank BenFord tested the aforementioned information on 20 different data sets and found it to be satisfying. Later on, BenFord was credited for his efforts.

2 Definition

BenFord's Law can be applied to a set of number if the first digit, δ appears in the set with a probability,

$$\pi(\delta) = log_{10}(\delta + 1) - log_{10}(\delta)$$
$$\pi(\delta) = log_{10}(1 + 1/\delta)$$

Note: $\delta \ \epsilon \ \{1, 2, 3, 4, ..., 9\}$

To get a better grasp, we should take one example, a good one would be the leading digit of 2^η.
One sequence from OEIS (On-Line Encyclopaedia of Integer Sequences) https://oeis.org/A008952 gives us the first digit of the sequence 2^η for the first 96 terms. After studying that sequence, we can get a table somewhat like,

δ	Count	Percentage	As per BenFord's Law
1	29	30.2%	30.1%
2	17	17.7%	17.6%
3	12	12.5%	12.5%
4	10	10.4%	9.7%
5	7	7.3%	7.9%
6	6	6.3%	6.7%
7	5	5.2%	5.8%
8	5	5.2%	5.1%
9	5	5.2%	4.6%

The digit 1 as the leading digit is the most probable situation.

Generalization to digits beyond the first It is possible to extend the law to digits beyond the first. Like first digit, the λ'th digit can also be taken into consideration for all $\lambda > 1$. The probability of occurrence of the digit, δ in the λ'th digit

3 Applications

BenFord's Law find it's importance at a wide range, such as

- Accounting fraud detection
- Criminal Trials
- Election Data
- Macroeconomic Data
- Price Digit Analysis
- Genome Data
- Scientific Fraud Detection