

Detecting Data Fraud ~Using Benford's Law~

KASHIMOTO Hiromori TOBA Eijin NAKASHIMA Haruka NISHIJIMA Hayaki MIZUNO Haruki

Abstract

Data fraud has been in the news a lot lately. In order to expose this data fraud, we have examined Benford's Law and created a standard for detecting fraud.

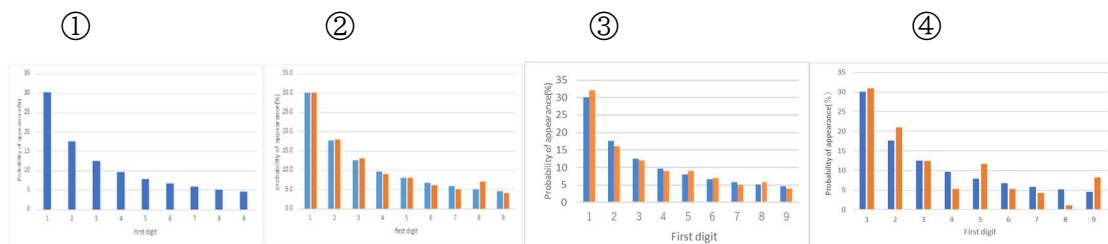
Introduction

Benford's law is a law that states that the leftmost digit of a number in nature has the highest number of 1s and decreases as the number goes up to 9. It is valid when the numbers are logarithmically distributed. The ultimate goal is to gain sufficient knowledge of Benford's Law to be able to use it in window dressing and accounting audits.

Theory and Experiment

Currently, after making a table using Excel to see what numbers are valid for Benford's Law, we are making graphs and recording those that come close to the ideal form of Benford's Law graph (①). For example, I have recorded graphs created by using a number sequence such as the Fibonacci sequence (②) and the equal ratio sequence (③), by taking the four-day total of the number of new coronavirus cases by prefecture in newspapers (④), as well as other graphs that were created during the research process. I am also in the process of recording the results of the research in which Benford's law was not established.

Translated with www.DeepL.com/Translator (free version)



* The blue graph is the ideal form of Benford's Law

Results

These were the "Equal Proportion Sequence" and the "Fibonacci Sequence" and the "randomly selected from newspapers" followed Benford's Law.

Discussion

Benford's Law does not hold for artificially allocated numbers or for those with fixed upper or lower numerical limits.

Conclusion

Our future vision is to use Benford's Law to create software that can detect fraud from a vast amount of data.

References

Newton April 2019, Wikipedia, "Benford's Law" applied by the economist who made Google the world, exposes fraud in bookkeeping, elections, etc., "Why Benford's Law of Mathematics?" and understand it intuitively! | Math Star

Keywords

Benford's Law, Fibonacci sequence, Equal ratio sequence, Number of coronavirus cases, Kobe Shimbun, First term, common ratio, number of items