Boise State University
ScholarWorks

MSS_2020 Mathematics Senior Showcase

12-1-2020

The History and Application of Benford's Law

Hunter Clark
Boise State University

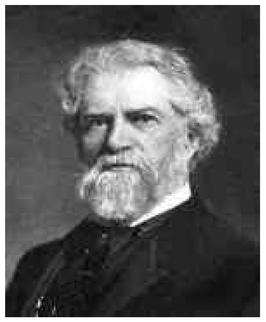
The History and Application of Benford's Law

Hunter Clark
Boise State University

BOISE STATE UNIVERSITY

Introduction

- Benford's Law is also known as the first digit law. It is a law that states the number 1 in a list or a set of data will show up approximately 30% more often then other numbers as a leading digit.
- In 1881 an Astronomer named Simon Newcomb, noticed a trend in his logarithm books. He noticed that the tables with smaller first digits were being used much more frequently
- Simon Newcomb was for the most part completely ignored until 1938 when Frank Benford decided to test Newcomb's theory. He ran a test consisting of 20,000 different data sets. His findings lead to a paper called "The Law of Anomalous Numbers."
- This paper is why credit was given to Benford for this discovery and is why the law is called Benford's law and not Newcomb's law.

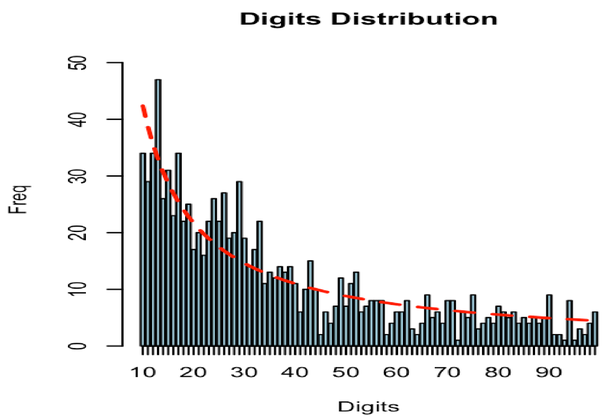


Simon Newcomb

Frank Benford

R-Code to Create a Benford plot (Data table was pulled from Dr. Ko's Class website under the data tab called BMW.dat)

- install.packages("benford.analysis")
  library(benford.analysis) # loads package
- Y <- scan("~/Downloads/BMW.dat")
- benf <- benford(y)
- plot(benf)



How Benford's Law is used today

- Benford's Law is widely used to assist in the detection of fraud and error among large sets of numbers such as taxes or book keeping.
- Benford's law usually applies to large sets of naturally occurring numbers. Some examples are stock market values, demographic data, tax data, logarithms, and scientific data.
- It is important to note that this law doesn't usually apply to data sets that have a minimum or maximum value. If numbers are assigned such as hourly rates, zip codes, or telephone numbers the set of numbers will not fit the law. This is why it is described as the Law of Anomalous Numbers.

How it works

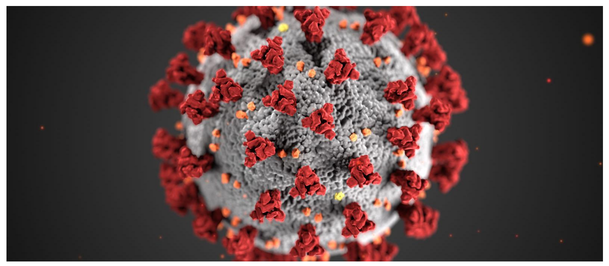
To identify fraud, an investigator would tally op the number of leading digits in financial statements.

The investigator would then compare the transactions to the expected numbers from Benford's Law.

Once compared, any transaction that does not fall within in the curve can be considered an outlier and should further be investigated.

Benford's use against COVID

- Benford's law is currently being used to determine if intervention and control is working to flatten the curve and slow down COVID.
- Because the number of infections are following an exponential trend, it is proven that exponential trends will automatically follow Benford's Law.
- Study's are taking the exponential data and comparing it against Benford's Law. Once intervention and controls are put into place they are then measuring again to see if it stop's following Benford's law. If it does stop following Benford's law this means the controls are working and the curve is being flattened.



Conclusion

- Simon Newcomb's observation and Frank Benford's curiosity unlocked another mathematical phenomenon that can likely be placed in the same category as Fibonacci number's and Pascal's triangle. These are things that show up time and time again and still to this day are still not fully understood.

Acknowledgements

- Simon Newcomb image- https://mathshistory.st-andrews.ac.uk/Biographies/Newcomb/
- Frank Benford image- https://alchetron.com/Frank-Benford
- COVID and Benford- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7431331/
- COVID image - https://www.fda.gov/food/food-safety-during-emergencies/food-safety-and-coronavirus-disease-2019-covid-19