Investigating Novel Analysis by Benford Methodology for Fraud and Anomaly Detection

Raju, Anish

Faked company payments and fraudulent schemes such as Ponzi schemes steal over $3.7 trillion from innocent consumers every year, with auditing methods significantly lagging behind the increase in fraudsters. Checking the accuracy of models of numerous natural processes, such as population growth and rural flooding, and of large databases relied upon by government sustainability programs, is also currently extremely difficult. The goal of this research project was to create a new way of analyzing datasets for manipulation or accuracy, by using leading digit probabilities predicted by Benford's Law. Conventional methods of analysis using Benford's Law tend to often overlook or understate discrepancies. In this experiment, a new, improved method for measuring conformity to Benford's Law is provided using mathematical tests based on ranges of variance from a model, derived logarithmic patterns, and cumulative probability distribution functions of leading digits. Examining dataset distributions after initially unsuccessful attempts led to the surprising discovery that many of the datasets followed an extreme variation of Benford's Law. Benford's Law was then expanded by adding a theorized correcting factor for functional relationships, which adjusts the predicted Benford proportions of first digits to account for how skewed a dataset is. Together, these tests enable a more complete, accurate analysis. Various types of datasets were analyzed using the new methodology, and criteria for which datasets tend to Benford's Law were also devised. Ultimately, the new method was significantly more effective than the Conventional method at detecting fraud and anomalies - almost a third of the datasets would not have been successfully analyzed if the new method had not been utilized.