Application of resampling methods when testing agreement with Benford's distribution

Vesna Rajić

University of Belgrade, Faculty of Economics and Business
e-mail: vesna.rajic@ekof.bg.ac.rs

Jelena Stanojević

University of Belgrade, Faculty of Economics and Business
e-mail: jelena.stanojevic@ekof.bg.ac.rs

Abstract. Benfords law is a statistical methodology that quickly and efficiently locates suspicious positions and figures in large dataset. This law can be used to detect anomalies, errors or frauds in different data sets, especially in financial reports of companies. Benfords law claims that leading digits from 1 to 9 appear in a decreasing logarithmic law. It means that the digit 1 appears the most frequently, followed by two, three, etc. The digit 9 has the smallest frequency according to this law. In this paper we investigate the conformity of an empirical distribution of observed real data with the Benford's distribution. In addition to the tests that are presented in the literature (z-test, chi-square test, Kolmogorov-Smirnov test, MAD test) the application of resampling methods is suggested. The main goal of the paper is to present an application of the bootstrap tests as well as the permutation tests for checking conformity with Benford's law and to indicate their comparative advantages and disadvantages.

Keywords: Benfords law, Data manipulation, Statistical tests, Resampling methods

References

[1] **F. Benford** The law of anomalous numbers. *Proceedings of the American philosophical society*, 1938, 551-572.
[2] **G. Fang, Q. Chen, Q.** Several common probability distributions obey Benfords law. *Physica A: Statistical Mechanics and its Applications*, 2020, 540, 123-129.
[3] **T. P. Hill** The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *merican Scientist*, 1998, 86(4), 358-363.
[4] **M. J. Nigrini** Benford's Law: Applications for forensic accounting, auditing, and fraud detection (Vol. 586). *John Wiley & Sons*, 2012.
[5] **W. A. Wallace** Assessing the quality of data used for benchmarking and decision-making. *The Journal of Government Financial Management*, 2002, 51(3), 16.