Two Digit Testing for Benford's Law

Dieter W. Joenssen*
University of Technology Ilmenau, Ilmenau, Germany
Dieter.Joenssen@TU-Ilmenau.de

Benford's law has been used by auditors to help reveal data manipulation not only in the context of tax audits and corporate accounting, but also election fraud. The principle idea behind Benford's law is that the frequency distribution of the first digits from numerical data of certain processes should conform to a discrete distribution known as the Benford distribution. Goodness-of-fit tests have been used to assess if the data's first digits conform to said distribution. When data should conform to Benford's law, a null-hypothesis rejection suggests that some form of data manipulation has taken place. Goodness-of-fit tests, like all tests of statistical significance, are prone not only to the type I error, which is limited by the chosen level of significance, but also to the type II error which decreases not only with sample size but is also inherently lower for some testing procedures than others. One possible procedural change is not to test the distribution of the data's first digit, as is the current standard, but to test the joint distribution of the data's first two digits. The gain in power would be due to an increased utilization of information, because, given that the null hypothesis is true, the distributions of the first and second digits are not independent. This paper describes how four goodness-of-fit tests can be extended to test the joint distribution of the first and second digit for conformity to Benford's law. Additionally, a comparison of power yielded by the original (one-digit) as well as the proposed (two-digit) analysis is provided.

Key Words: fraud detection, goodness-of-fit test, power comparison, type II error