



Usability of the Benford's Law for Time Series

Nursu Tunalioglu and Bahattin Erdogan

Yildiz Technical University, Civil Engineering Faculty, Department of Geomatic Engineering, Istanbul, Turkey
(ntunali@yildiz.edu.tr)

Benford's law (BL), also known as the First-digit or Significant-digit law, is an intriguing pattern in data sets, considers the frequency of occurrence of the first digits, which are not uniformly distributed as might be expected, conversely follow a specified theoretical -logarithmic- distribution. In BL, the first leading (i.e. first non-zero) decimal digit is not equally likely to be any one of the nine possible digits. Instead of that, the occurrence of numbers starting with 1 and 2 is close to 30% and 18%, respectively, whereas numbers starting with 8 or 9 are close to 5%. In this paper, we aim to demonstrate consistency of the BL in data sets assumed that contain uniform distribution and represent whether or not certain data sets follow BL. We use two groups of data sets: 1. simulated, 2. real (GNSS data). The first group data set is simulated based on linear regression and univariate models and each simulated group is generated for a number of 100, 1.000, and 10.000. To generate second group, an IGS station data (ISTA) is processed by real time PPP approach using BKG Ntrip Client (BNC v2.11.2) software. Here, the observation duration of GNSS data is 3 hours. The first 30 minutes of observation duration is removed from the data set on analyzes according to problems on ambiguity resolution. BL has been applied to the residuals (v) and normalized residuals (w) estimated from data sets. Goodness-of-fit test, which is based on how good a fit we have between the frequency of occurrence of observations in an observed sample and the expected frequencies obtained from the hypothesized distribution, has been implemented to determine if a population has a specified BL distribution. The results show that each data set used in this study follows disconcertingly BL.

Keywords: Benford' Law, First Digit Phenomena, Frequency of Occurrence, Occurrence of observations, Goodness-of-fit test