An Integrated Approach to Electoral Analysis: Benford's Law, Bayes Factors and Electoral Footprints

THESIS
DEPARTMENT OF
MATHEMATICS

Daniel Anglada Marrero,
daniel.anglada@upr.edu
Department of Mathematics, UPR-RP.

May 11, 2026

Abstract

This thesis presents an integrated statistical approach for the forensic analysis of electoral data by combining Benford's Law, Bayes factors, and a geometric representation of voting patterns known as the electoral footprint. The study focuses on the exploratory identification of atypical patterns or statistical deviations in election results using probabilistic tools and statistical models, rather than relying solely on descriptive methods.

First, the thesis examines whether electoral data follow Benford's Law by analyzing the distribution of digits and performing goodness-of-fit tests. Based on these results, Bayesian models are compared using Bayes factors to evaluate competing hypotheses about the mechanisms that generate the data. This approach provides a clearer and more direct measure of the evidence than traditional hypothesis testing in this context.

In addition, the thesis introduces the electoral footprint, which represents the relationship between voter turnout and the proportion of votes received across polling stations. This tool makes it possible to detect patterns in the data that are not readily apparent through simple inspection. By using this representation, it is possible to identify atypical behaviors that deviate from the expected patterns.

The proposed methodology is applied to electoral data from Puerto Rico for the 2020 and 2024 elections. The results illustrate how the combination of Benford's Law, Bayes factors, and the electoral footprint improves the analysis of electoral data and facilitates the interpretation of potential anomalies. Overall, the findings highlight the usefulness of Bayesian methods in this setting and demonstrate that this approach provides a valuable methodological framework for complementing the exploratory analysis of electoral data.