Multiple Round Ballot Polling Risk-Limiting Audit Simulations

Abstract

Risk-Limiting Audits (RLAs) guarantee with known probability that if the outcome of an election is incorrectly announced, it will be detected, and a full hand recount will be performed. In Ballot-polling RLAs, samples of ballots are drawn and tallied. In this paper we present an audit simulation framework that can be used to confirm theoretical results of audits. We present such experimental results for multiple round ballot-polling RLAs. BRAVO [citation] has long been the standard ballot-polling RLA, while Minerva was recently introduced with the claim that fewer ballots on average are necessary to conclude an audit. [Minerva paper] presented experimental results for one round audits, and here we present results from simulations of multiple round audits using Minerva as well as BRAVO (both Selection-Ordered BRAVO and End-of-Round BRAVO). The simulation results agree within reasonable error with the mathematical properties claimed in [Minerva paper]. On average, BRAVO audits are unnecessarily conservative while Minerva audits stop with fewer ballots. We also present details on software implementing Minerva and the simulations themselves.

1 Introduction: B2 and R2 Ballot Polling RLAs

In ballot-polling RLAs, samples of ballots are drawn and tallied in rounds after each of which a statistical measure determines whether to continue.

2 Simulations to Confirm Theoretical Results

The outcomes of RLAs depend on random chance; some random samples support the alternative hypothesis more than expected, resulting in quick low-risk conclusions, while other unlucky samples require subsequent rounds to confirm the announced result. We can simulate random samples for various true underlying ballot distributions by computing pseudorandom samples. By applying an audit's stopping condition to thousands of such simulated samples, the stopping behavior will tend towards the true frequencies for the audit. In this way, we can examine whether theoretical claims for an audit are actually correct.

For this paper, we simulated audits for all US states with at least a 5% margin in the 2020 Presidential election. For each of these states, we simulated $10,000 = 10^4$ audits with the announced underyling ballot distribution and an additional $10,000 = 10^4$ audits with a tie as the underlying ballot distribution.

2.1 End-of-Round BRAVO

2.2 Selection-Ordered BRAVO

2.3 Minerva