On Estimating the Size and Confidence of a Statistical Audit

Javed A. Aslam
College of Computer and Information Science
Northeastern University

Raluca A. Popa and Ronald L. Rivest
Computer Science and Artificial Intelligence Laboratory
M.I.T.

\[ u = \left( n - \frac{b - 1}{2} \right) \cdot (1 - (1 - c)^{1/b}) \]
Outline

- Motivation
- Background
  - How Do We Audit?
  - The Problem
- Analysis
  - Model
  - Sample Size
  - Bounds
- Conclusions
Motivation

- There have been cases of electoral fraud (Gumbel’s *Steal This Vote*, Nation Books, 2005)
- Would like to ensure confidence in elections
- Auditing = comparing statistical sample of paper ballots to electronic tally
- Provides confidence in a software independent manner
How Do We Audit?

- Proposed Legislation: *Holt Bill (2007)*
  - Voter-verified paper ballots
  - Manual auditing
- Granularity: Machine, Precinct, County
- Procedure
  - Determine $u$, # precincts to audit, from margin of victory
  - Sample $u$ precincts randomly
  - Compare hand count of paper ballots to electronic tally in sampled precincts
    - If all are sufficiently close, declare electronic result final
    - If any are significantly different, investigate!
How Do We Audit?

- Proposed Legislation: *Holt Bill (2007)*
  - Voter-verified paper ballots
  - Manual auditing
- Granularity: Machine, Precinct, County
- Procedure
  - Determine $u$, # precincts to audit, from margin of victory
  - Sample $u$ precincts randomly
  - Compare hand count of paper ballots to electronic tally in sampled precincts

Our formulas are independent of the auditing procedure
The Problem

- How many precincts should one audit to ensure high confidence in an election result?
Previous Work

- *Saltman (1975)*: The first to study auditing by sampling without replacement
- *Dopp and Stenger (2006)*: Choosing appropriate audit sizes
- *Alvarez et al. (2005)*: Study of real case auditing of punch-card machines
Hypothesis Testing

- Null hypothesis: The reported election outcome is incorrect (electronic tally indicates different winner than paper ballots)
- Want to reject the null hypothesis
  - Need to sample enough precincts to ensure that, if no fraud is detected, the election outcome is correct with high confidence
Model

$n$ precincts  ○  ○  ●  ○  ●  ○  $b$ corrupted ("bad")

Sample $u$ precincts  ○  ○  ●
(without replacement)

- $c =$ desired confidence
- Want: If there are $\geq b$ corrupted precincts, then sample contains at least one with probability $\geq c$
- **Equivalently:** If the sample contains no corrupted precincts, then the election outcome is correct with probability $\geq c$
- Typical values: $n = 400$, $b = 50$, $c = 95\%$
What is $b$?

- Minimum # of precincts adversary must corrupt to change election outcome
- Derived from margin of victory

\[ b = \text{(half margin of victory)} \cdot n \]

- Our formulas are independent of $b$’s calculation

(times 5 (Dopp and Stenger, 2006))
Rule of Three

- If we draw a sample of size $\geq 3n/b$ with replacement, then:
  - Expect to see at least three corrupted precincts
  - Will see at least one corrupted precinct with $c \geq 95$

- In practice, we sample without replacement (no repeated precincts)
Sample Size

- Probability that no corrupted precinct is detected:
  \[ Pr = \binom{n-b}{u} / \binom{n}{u} \]
- Optimal Sample Size: Minimum \( u \) such that \( Pr \leq 1 - c \)

\( \rightarrow \) Problem: Need a computer

- Goal: Derive a simple and accurate upper bound that an election official can compute on a hand-held calculator
Our Bounds

- Intuition: How many different precincts are sampled by the Rule of Three?
- Our without replacement upper bounds:

\[ u = \left\lfloor n \cdot (1 - (1 - c)^{1/b}) \right\rfloor \]

\[ u = \left\lfloor (n - \frac{b - 1}{2}) \cdot (1 - (1 - c)^{1/b}) \right\rfloor \]

\[ u = \left\lfloor \frac{b}{H_n - H_{n-b}} \cdot (1 - (1 - c)^{1/b}) \right\rfloor \]
Our Bounds

- Intuition: How many different precincts are sampled by the Rule of Three?
- Our *without replacement* upper bounds:

\[
u = \left\lceil n \cdot (1 - (1 - c)^{1/b}) \right\rceil \\
u = \left\lceil \left( n - \frac{b - 1}{2} \right) \cdot (1 - (1 - c)^{1/b}) \right\rceil
\]

- Example: \( n = 400, \ b = 50 \) (margin=5%), \( c = 95\% \)

\[
u = \left\lceil \left( 400 - \frac{50 - 1}{2} \right) \cdot (1 - (1 - 0.95)^{1/50}) \right\rceil = \left\lceil 21.84 \right\rceil = 22
\]
Our Bound

\[ u = \left[ \left( n - \frac{b - 1}{2} \right) \cdot (1 - (1 - c)^{1/b}) \right] \]

- Conservative: provably an upper bound
- Accurate:
  - For \( n \leq 10,000, b \leq n/2, c \leq 0.99 \) (steps of 0.01):
    - 99% is exact, 1% overestimates by 1 precinct
    - Analytically, it overestimates by at most \(-\ln(1-c)/2\), e.g. three precincts for \( c < 0.9975 \)
- Can be computed on a hand-held calculator

\[(1 - c)^{1/b} = e^{\ln(1-c)/b}\]
Observations

Precincts to Audit

- Fixed level of auditing is not appropriate

n = 400, c=95%

Optimal Sample Size
Our Bound

Margin of Victory

- 0%
- 1%
- 5%
- 10%
- 15%
- 20%
- 25%
Observations (cont’d)

Precincts to Audit

- **Holt Bill (2007):** Tiered auditing
Related Problems

- Inverse questions
  - Estimate confidence level $c$ from $u$, $b$, and $n$
  - Estimate detectable fraud level $b$ from $u$, $c$, and $n$
- Auditing with constraints
  - *Holt Bill (2007):* Audit at least one precinct in each county
- Future work
  - Handling precincts of variable sizes (*Stanislevic, 2006*)
Conclusions

- We develop a formula for the sample size:

\[
    u = \left( n - \frac{b - 1}{2} \right) \cdot (1 - (1 - c)^{1/b})
\]

that is:
- Conservative (an upper bound)
- Accurate
- Simple, easy to compute on a pocket calculator
- Applicable to different other settings
Thank you!

- Questions?