Risk-limiting Audits and Evidence-based Elections in a Nutshell

Philip B. Stark

Department of Statistics
University of California, Berkeley

EVN Annual Conference
Atlanta, GA
14 March 2013

What do we want election audits to do?

Ensure that the electoral outcome is correct;
If outcome is wrong, correct it before it's official.

How can an audit correct a wrong outcome?

If there's an adequately accurate audit trail, the audit could in principle count all the votes by hand.

Don't overturn outcome on statistics alone.

Why not just count all votes by hand?

- Unnecessarily expensive and time-consuming.
- Instead, count by machine, and check a random sample by hand.
- Keep checking until there's convincing evidence that the outcome is right—or until all ballots have been examined and the right outcome is known.

Controlling the chance of error

- Since the sample is drawn at random, there's a chance a wrong outcome will escape correction—but we can make that chance as small as we want. Statistics says how.
- *Risk* is the largest possible chance that the audit does not correct the outcome, if the outcome is wrong.
- *Risk-limiting audit* ensures that the largest possible chance is still a small chance, like 10%, 5%, 1%.
- Generally, have to check more to make chance smaller.

Random Sampling

"Stirring" is key to reducing work

- Don't have to climb into the bathtub to tell if it's hot: can just stick your toe in—if the water is stirred well.
- Don't have to walk all over town to tell if it's cold outside: the air is mixed well enough that you just have to step outside to get a pretty good idea.
- Don't have to drink a whole pot of soup to tell if it's too salty: a teaspoon is enough—if the pot has been stirred. (Doesn't matter whether the pot holds 1q or 50g.)

How do you stir ballots?

Random sampling is stirring

- Imagine numbering the ballots.
- Write the numbers on ping-pong balls; put in a lotto machine.
- Lotto machine stirs the balls and spits some out.
- The ballots with the numbers on the selected balls are a random sample of ballots.
- Easier to stir balls than ballots. Even easier to generate random numbers.
- Still amounts to putting ballots into a huge cement mixer to stir them, then taking a "teaspoon" of ballots.

Risk is *not*

- The chance that the certified outcome is wrong.
- The fraction of certified outcomes that are wrong.

Paper rules—if it is right

- Can't correct wrong outcomes without counting the whole audit trail.
- Counting the whole audit trail won't give right answer unless it's adequately accurate and intact.
- Current procedures for protecting, tracking, and accounting for ballots are spotty. Should be top priority!

Risk limit *assumes* outcome is wrong in the hardest-to-find way. Biggest chance the outcome won't be corrected.

Ballot-polling Audits and Comparison Audits

- Ballot polling audit: sample ballots until there is strong evidence that looking at all of them would show the same election outcome.
  Like an exit poll—but of ballots, not voters.
- Comparison audit:
  1. Commit to vote subtotals (or CVRs), e.g., precinct-level results
  2. Check that the subtotals add up exactly to contest results
  3. Check subtotals by hand until there is strong evidence the outcome is right

Tradeoffs

- Ballot polling audit
  - Virtually no set-up costs
  - Requires nothing of voting system
  - Need a ballot manifest to draw sample
  - Preserves voter anonymity except possibly for sampled ballots
  - Requires more counting than ballot-level comparison audit
  - Does not check tabulation: outcome could be right because errors cancel
- Comparison audit
  - Heavy demands on voting system for reporting and data export
  - Requires LEO to commit to subtotals
  - Requires ability to retrieve ballots that correspond to CVRs or subtotals
  - May compromise voter privacy
  - Most efficient (ballot-level) not possible w/ current systems: requires rescan
  - Checks tabulation (but not for *transitive audits* unless subtotals are cross checked as well)
  - Ballot-level comparison audits require least hand counting

Ballot-polling Audits are often Cheap for Big Contests

255 state-level presidential contests, 1992–2011, 10% risk limit

BPA expected to examine fewer than 308 ballots for half the contests.

Work expands as margins shrink, but we could get a lot of election integrity at low cost—with any paper-based system.

Workload estimate: Ballot-Polling Audit, 2 Candidates, 10% Risk Limit

Winner's True Share	Ballots drawn		
	median	90th percentile	Mean
70%	22	60	30
65%	38	108	53
60%	84	244	119
58%	131	381	184
55%	332	974	469
54%	518	1,520	730
53%	914	2,700	1,294
52%	2,051	6,053	2,900
51%	8,157	24,149	11,556
50.5%	32,547	96,411	46,126

Risk-Limiting Audits

- 17 pilot audits in CA, CO, and OH; another 13 planned.
- CO has law; CA has pilot law
- simple measures, super-majority, multi-candidate, vote-for-n
- multiple contests audited simultaneously with one sample
- contest sizes: 200 ballots to 121,000 ballots
- counting burden: 16 ballots to 7,000 ballots
- cost per audited ballot: nil to about $0.55
- several jurisdictions have audited on their own—no geeks needed

Evidence-based elections

Principle: Trust, but verify

LEOs should give convincing evidence that outcomes are right (or say they can't).
"Trust me" is not convincing.

- Voters create complete, durable, accurate audit trail.
- LEO curates the audit trail adequately.
- Compliance audit to check whether the audit trail is trustworthy enough to determine who won.
  If not, how strong can the evidence be?
- Risk-limiting audit to correct the outcome if it is wrong.
  Presumes audit trail is OK.
  "Explaining" or "resolving" errors isn't enough.