Bayesian Post-Election Audits

Ronald L. Rivest and Emily Shen

Viterbi Professor of EECS
MIT, Cambridge, MA
{rivest,ehshen}@mit.edu

EVT/WOTE 2012
2012-08-07

Outline

Post-Election Audit Objectives

By examining by hand sufficiently many randomly selected paper ballots:

- Confirm to a high degree of confidence that the reported (scanner-based) outcome is correct or else that the actual (full hand-count) outcome is different.

Post-Election Audit Objectives

By examining by hand sufficiently many randomly selected paper ballots:

- Confirm to a high degree of confidence that the reported (scanner-based) outcome is correct or else that the actual (full hand-count) outcome is different.
- Convince the losers they really lost!

Single-ballot Audits

- Sequential decision-making (Wald).
- Examine paper ballots *one at a time*, in random order.
- Determine *actual type* of each ballot (as opposed to its *reported type*).
- At each stage, decide whether to
  - Stop: Reported outcome looks **OK**.
  - Continue: more auditing needed.

  (We assume that full hand count needed to overturn reported outcome.)

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand):

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): ? ? ? ? ?

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A ? ? ? ?

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A A ? ? ?

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A A ? ? ?

- *Comparison audit:* also look at their *reported types:*

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A A ? ? ?

- *Comparison audit:* also look at their *reported types:*

  reported types (scanner): A A A B A

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A A ? ? ?

- *Comparison audit:* also look at their *reported types:*

  reported types (scanner): A A A B A
  actual types (by hand): ? ? ? ? ?

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A A ? ? ?

- *Comparison audit:* also look at their *reported types:*

  reported types (scanner): A A A B A
  actual types (by hand): A ? ? ? ?

Ballot-polling audits vs. comparison audits

- *Ballot-polling audit:* look at only the *actual ballot types* of the audited ballots:

  actual types (by hand): A A ? ? ?

- *Comparison audit:* also look at their *reported types:*

  reported types (scanner): A A A B A
  actual types (by hand): A A ? ? ?

Auditing with a magic box

- Assume you audit randomly chosen ballots, one by one, in a ballot-polling audit.

Auditing with a magic box

- Assume you audit randomly chosen ballots, one by one, in a ballot-polling audit.
- Suppose I give you a “magic box” that at any time can answer the question,

Auditing with a magic box

- Assume you audit randomly chosen ballots, one by one, in a ballot-polling audit.
- Suppose I give you a "magic box" that at any time can answer the question,

  Given what you've seen in the audit so far, what is the probability that each candidate would win if all ballots were examined?

Auditing with a magic box

- Assume you audit randomly chosen ballots, one by one, in a ballot-polling audit.
- Suppose I give you a "magic box" that at any time can answer the question,

  Given what you've seen in the audit so far, what is the probability that each candidate would win if all ballots were examined?

- Then you can stop audit if/when the reported winner has at least (say) 95% probability of winning.

Actual ballot types (by hand):

? ? ? ? ? ? ? ? ? ? . . .

Probability A wins: 50.0%

Probability B wins: 50.0%

An example

Actual ballot types (by hand):

A ? ? ? ? ? ? ? ? ?…

Probability A wins: 75.0%
Probability B wins: 25.0%

An example

Actual ballot types (by hand):

A A ? ? ? ? ? ? ? ? . . .

Probability A wins: 87.5%

Probability B wins: 12.5%

An example

Actual ballot types (by hand):

A A B ? ? ? ? ? ? ? . . .

Probability A wins: 68.8%

Probability B wins: 31.2%

An example

Actual ballot types (by hand):

A A B B ? ? ? ? ? ? . . .

Probability A wins: 50.0%

Probability B wins: 50.0%

An example

Actual ballot types (by hand):

A A B B A ? ? ? ? ? . . .

Probability A wins: 65.6%

Probability B wins: 34.4%

An example

Actual ballot types (by hand):

A A B B A A ? ? ? ? . . .

Probability A wins: 77.4%

Probability B wins: 22.6%

An example

Actual ballot types (by hand):

A A B B A A A ? ? ? . . .

Probability A wins: 85.6%

Probability B wins: 14.4%

An example

Actual ballot types (by hand):

A A B B A A A A ? ? . . .

Probability A wins: 91.0%

Probability B wins: 9.0%

An example

Actual ballot types (by hand):

A A B B A A A A A ? . . .

Probability A wins: 94.5%

Probability B wins: 5.5%

An example

Actual ballot types (by hand):

A A B B A A A A A A . . .

Probability A wins: 96.7%

Probability B wins: 3.3%

An example

Actual ballot types (by hand):

A A B B A A A A A A . . .

Probability A wins: 96.7%

Probability B wins: 3.3%

$\rightarrow$ Stop auditing! $\leftarrow$

Making the magic box (ballot-polling)

- Suppose you are auditing an election between candidates A and B, with 5 ballots.

Making the magic box (ballot-polling)

- Suppose you are auditing an election between candidates A and B, with 5 ballots.
- You draw a random sample (without replacement) of two ballots.

Making the magic box (ballot-polling)

- Suppose you are auditing an election between candidates A and B, with 5 ballots.
- You draw a random sample (without replacement) of two ballots.
- Both ballots are for A:

A ? A ? ?

Making the magic box (ballot-polling)

- Suppose you are auditing an election between candidates A and B, with 5 ballots.
- You draw a random sample (without replacement) of two ballots.
- Both ballots are for A:

A ? A ? ?

- *Q: What is the probability that A won?*

Answer (Bayesian)

- To make Q well-posed, need a model (a *prior*) for the likelihood of different outcomes.

Answer (Bayesian)

- To make Q well-posed, need a model (a *prior*) for the likelihood of different outcomes.
- A *noninformative prior* gives each outcome (A:B tally) equal probability:

tally	5:0	4:1	3:2	2:3	1:4	0:5
Prob	1/6	1/6	1/6	1/6	1/6	1/6

Answer (Bayesian)

- To make Q well-posed, need a model (a *prior*) for the likelihood of different outcomes.
- A *noninformative prior* gives each outcome (A:B tally) equal probability:

tally	5:0	4:1	3:2	2:3	1:4	0:5
Prob	1/6	1/6	1/6	1/6	1/6	1/6

- With this prior and sample, A wins with (subjective) probability

95%

Answer (Bayesian)

- To make Q well-posed, need a model (a *prior*) for the likelihood of different outcomes.
- A *noninformative prior* gives each outcome (A:B tally) equal probability:

tally	5:0	4:1	3:2	2:3	1:4	0:5
Prob	1/6	1/6	1/6	1/6	1/6	1/6

- With this prior and sample, A wins with (subjective) probability

95%

- If your error limit is 5%, stop auditing!

95%? (Bayes Rule)

posterior probability proportional to:

- prior times likelihood of sample given prior

95%? (Bayes Rule)

posterior probability proportional to:

prior times likelihood of sample given prior

tally	5:0	4:1	3:2	2:3	1:4	0:5
prior	1/6	1/6	1/6	1/6	1/6	1/6

95%? (Bayes Rule)

posterior probability proportional to:

prior times likelihood of sample given prior

tally	5:0	4:1	3:2	2:3	1:4	0:5
prior	1/6	1/6	1/6	1/6	1/6	1/6
likelihood(AA)	$\frac{5}{5} \cdot \frac{4}{4}$	$\frac{4}{5} \cdot \frac{3}{4}$	$\frac{3}{5} \cdot \frac{2}{4}$	$\frac{2}{5} \cdot \frac{1}{4}$	0	0

95%? (Bayes Rule)

posterior probability proportional to:

prior times likelihood of sample given prior

tally	5:0	4:1	3:2	2:3	1:4	0:5
prior	1/6	1/6	1/6	1/6	1/6	1/6
likelihood(AA)	$\frac{5}{5} \cdot \frac{4}{4}$	$\frac{4}{5} \cdot \frac{3}{4}$	$\frac{3}{5} \cdot \frac{2}{4}$	$\frac{2}{5} \cdot \frac{1}{4}$	0	0
product	$\frac{10}{60}$	$\frac{6}{60}$	$\frac{3}{60}$	$\frac{1}{60}$	0	0

95%? (Bayes Rule)

posterior probability proportional to:

prior times likelihood of sample given prior

tally	5:0	4:1	3:2	2:3	1:4	0:5
prior	1/6	1/6	1/6	1/6	1/6	1/6
likelihood(AA)	$\frac{5}{5} \cdot \frac{4}{4}$	$\frac{4}{5} \cdot \frac{3}{4}$	$\frac{3}{5} \cdot \frac{2}{4}$	$\frac{2}{5} \cdot \frac{1}{4}$	0	0
product	$\frac{10}{60}$	$\frac{6}{60}$	$\frac{3}{60}$	$\frac{1}{60}$	0	0
posterior	$\frac{10}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	0	0

95%? (Bayes Rule)

posterior probability proportional to:

prior times likelihood of sample given prior

tally	5:0	4:1	3:2	2:3	1:4	0:5
prior	1/6	1/6	1/6	1/6	1/6	1/6
likelihood(AA)	$\frac{5}{5} \cdot \frac{4}{4}$	$\frac{4}{5} \cdot \frac{3}{4}$	$\frac{3}{5} \cdot \frac{2}{4}$	$\frac{2}{5} \cdot \frac{1}{4}$	0	0
product	$\frac{10}{60}$	$\frac{6}{60}$	$\frac{3}{60}$	$\frac{1}{60}$	0	0
posterior	$\frac{10}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{1}{20}$	0	0
	A wins 95%			B wins 5%		

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn

Urn = {   }

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls

Urn = { }

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls

Urn = { A }

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls.

   Urn = { A }

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls.

$$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \qquad\qquad\qquad\qquad \}$$

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

$$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \qquad\qquad\qquad \}$$

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   Urn = $\{ \underbrace{A\,B}_{prior} \qquad\qquad\qquad\qquad \}$
3. Add s balls for sample (e.g. A A)

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   Urn = { $\underbrace{\text{A B}}_{\textit{prior}}$ A }
3. Add s balls for sample (e.g. A A)

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \quad \underbrace{A\,A}_{\textit{sample}} \qquad\qquad\qquad \}$$

3. Add s balls for sample (e.g. A A)

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \quad \underbrace{A\,A}_{\textit{sample}} \qquad\qquad \}$$

3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \quad \underbrace{A\,A}_{\textit{sample}} \quad A \qquad \}$$

3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{\ \underbrace{A\,B}_{\textit{prior}}\ \ \underbrace{A\,A}_{\textit{sample}}\ \ A\ \ B\qquad\}$$

3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \quad \underbrace{A\,A}_{\textit{sample}} \quad A \;\; B \;\; A \}$$

3. Add s balls for sample (e.g. AA)
4. $n - s$ times: duplicate a random ball in urn.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \underbrace{A\,B}_{\textit{prior}} \ \underbrace{A\,A}_{\textit{sample}} \ A \ B \ A \ \}$$

3. Add s balls for sample (e.g. AA)
4. $n - s$ times: duplicate a random ball in urn.
5. Remove balls added in step 2.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \quad \underbrace{\text{A A}}_{\textit{sample}} \quad \text{A} \quad \text{B} \quad \text{A} \quad \}$$

3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.
5. Remove balls added in step 2.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \quad \underbrace{A\,A}_{\textit{sample}} \quad A \;\; B \;\; A \; \}$$
3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.
5. Remove balls added in step 2.
6. Tally the n balls in the urn.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   Urn = { $\underbrace{\text{A A}}_{\textit{sample}}$ A B A } $\Longrightarrow$ tally 4:1
3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.
5. Remove balls added in step 2.
6. Tally the n balls in the urn.

Efficient sampling of posterior using Polya's Urn

1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   Urn = { $\underbrace{\text{A A}}_{\textit{sample}}$ A B A } $\Longrightarrow$ tally 4:1
3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.
5. Remove balls added in step 2.
6. Tally the n balls in the urn.

   Tally has desired distribution! (e.g. $\frac{10}{20}, \frac{6}{20}, \frac{3}{20}, \frac{1}{20}$)

Efficient sampling of posterior using Polya's Urn

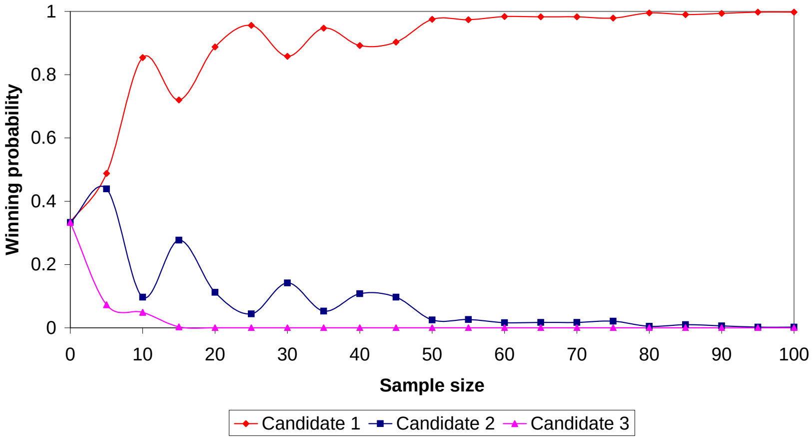
1. Given sample of size s from n-ballot profile.
2. In an urn put $a = 1$ "A" balls and $b = 1$ "B" balls. (a, b are *hyperparameters* for prior)

   $$\text{Urn} = \{ \underbrace{A\,A}_{sample} \;\; A \;\; B \;\; A \;\} \Longrightarrow \text{tally } 4{:}1$$
3. Add s balls for sample (e.g. A A)
4. $n - s$ times: duplicate a random ball in urn.
5. Remove balls added in step 2.
6. Tally the n balls in the urn.

   Tally has desired distribution! (e.g. $\frac{10}{20}, \frac{6}{20}, \frac{3}{20}, \frac{1}{20}$)

Can sample faster using gamma variates (see paper).

Winning probabilities vs. sample size in a Bayes audit



Arbitrary voting system

We note that a Bayes audit works for an *arbitrary voting system* as long as the number of ballot types is not too large; all you need is a way to compute the winner of a profile of ballots, and a way of sampling ballots. We have tested it on

- plurality
- IRV
- Borda
- Schulze

with good results.

Bayesian comparison audits

- Same idea, but have *one urn for each reported type*.
- Much more efficient!! (But needs way of matching paper ballots with their reported types.)

2011 Monterey Peninsula Water Mgt District Director

- Ballot-polling.
- Two candidates (plus write-ins).
- 2011 votes cast: 1353 for Lewis, 742 for Mancini (reported).
- Stark's ballot-polling audit with 10% risk limit examined:

  89 ballots.

- A Bayes ballot-polling audit with $\epsilon = 0.10$ examines:

  23 ballots on average

  11 ballots (median)

2011 Stanislaus Oakdale Measure O

- Comparison audit.
- Yes/No proposition.
- 3152 votes cast: 1728 Yes, 1392 No, 32 undervotes (reported).
- Stark's comparison audit with 10% risk limit examined:

  49 ballots.
- A Bayes ballot-polling audit with $\epsilon = 0.10$ examines:

  92 ballots (average)

  39 ballots (median).

Discussion

- We conjecture that a Bayes audit is in fact "risk-limiting" (perhaps given some suitable assumptions or constant factors in parameterization). But this is just a conjecture.
- The Bayes audit admits the use of other priors, such as those a very partisan observer might have.
- The Bayes audit admits the use of *multiple* priors; only stopping when all auditors (with different priors) agree to do so.

Summary – Bayes Audit Advantages

- High efficiency (few ballots get audited).
- Small/controllable miscertification rates observed.
- Simple in structure / easy to implement.
- Handles ballot-polling audits, comparison audits, and many different voting systems.
- No MOV computation required to start.
- Admits flexible (multiple) choice(s) of prior.
- Can be stopped early with meaningful results.

Summary – Bayes Audit Disadvantages

- Only works (so far) for single-ballot audits.
- Unclear relationship to risk-limiting audits.
- Results depend on choice(s) for prior.
- Need program to compute winning probabilities.

The End

For more info and code, contact authors or see:

`http://people.csail.mit.edu/rivest/bayes/`