A New Statistical Audit for Real Elections

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Grant McClearn & Sarah Morin

Neal McBurnett, Poorvi L. Vora, Filip Zagorski

GW | CS

A new risk-limiting audit for real elections, the Round-by-Round *RLA*

- implemented in software for use in pilots in this year's elections
- likelihood of being more widely used in 2020
- reduces average number of ballots examined by about 25% or more.

Background and Related Work

- Security recommendation: audit election paper trail to ensure the reported winner really won.
- Use a *risk-limiting audit* (*RLA*).
  - The *risk* of an audit is:

    Pr[election passes audit | election is incorrect]
  - Type I error in a binary hypothesis test, where the null hypothesis is that the election outcome is incorrect.
  - Risk-limiting audits guarantee the risk is lower than a pre-specified bound, the *risk limit*, independent of the underlying election
- If election does not pass audit, do a full hand count.
- Stark's *BRAVO RLA* [1]: used in governmental elections.
- Vora's Bayesian RLA [2], unifies the frameworks of *BRAVO* and Rivest's Bayesian audits [3]; the latter are not always *RLA*s.
- Existing theoretical audits assume:
  - Auditor samples ballot-by-ballot
  - Audit software determines whether the audit should stop after each ballot draw.
- In practice, however:
  - Ballots are drawn in rounds—say, 200 ballots, then 400, etc.
  - Audit software uses same rules as for ballot-by-ballot decisions, but takes them round-by-round.
  - Should use different rules because more information available at the time of the decision.
- But ballot-by-ballot stopping rules continue to be used.
- Current audits are too conservative: do not fully utilize the risk budget.

Notation

- N: total number of ballots cast in the two-candidate election.
- m: number of rounds of the audit.
- α: risk limit of the audit.
- r_j: number of invlaid ballots found in the audit sample for round j.
- $rnd(j)$: a "round schedule" function defined on the domain $\{j \in \mathbb{Z} : 1 \leq j \leq m\}$ which gives the number of ballots drawn by round j, inclusive.
- $rsk(j)$: a strictly positive "risk schedule" function defined on the domain $\{j \in \mathbb{Z} : 1 \leq j \leq m\}$ such that $\sum_j rsk(j) = \alpha$.
- $D_{b,j}(i)$: the probability of drawing i ballots for the reported winner in (i.e, during) round j, where b ballots for the winner have already been drawn (and so $\sum_i D_{b,j}(i) = 1, \forall b, j$).
- $C_j(i)$: the probability of having accumulated i ballots for the reported winner by round j, inclusive, given adherence to the stopping rule in previous rounds.

References

[1] M. Lindeman and P. B. Stark. A gentle introduction to risk-limiting audits. *IEEE Security & Privacy*, 10(5):42–49, 2012.

[2] P. L. Vora. Risk-Limiting Bayesian Polling Audits for Two Candidate Elections. 2019.

[3] R. L. Rivest and E. Shen. A Bayesian method for auditing elections. *EVT/WOTE*, 2012.

Procedure for the Round-by-Round Risk-Limiting Audit

1. Before the audit begins, determine a round schedule and a risk schedule.
2. Conduct the first round: draw the number of ballots dictated by the round schedule.
3. Check against the audit software, that is, if the auditor received ≥ k^+ ballots for the reported winner.
   a. If so, the audit stops, and the reported outcome is certified.
   b. If not, proceed to the next round and repeat steps 2. and 3. Do not replace the ballots that were drawn.
4. If all rounds have been completed and the audit did not stop, perform a full hand count of the ballots in the election. (Failure to do so means that the audit will not reach its stated risk limit.)

Average Number of Ballots Examined

Different risk schedules lead to different averages of number of ballots examined.

Average Number of Ballots Examined (Round Schedule: [200, 400, 800, 1600, 3200])

Audit Type / Margin	2%	5%	10%	20%	50%
BRAVO-like*	97,943	33,222	1,649	312	200
Round-by-Round[1]	74,180	16,939	912	316	200
Round-by-Round[2]	78,089	21,879	825	274	200
Round-by-Round[3]	84,186	34,878	848	249	200

* BRAVO-like stopping rules applied to round-by-round decisions.
1 Increasing geometrically (common ratio of 3/2) risk schedule.
2 Uniform risk schedule.
3 Decreasing geometrically (common ratio of 1/2) risk schedule.

Note: N: 100,000, Risk-limit: 5%

Computation of the Round-by-Round *RLA*

For the first round, there is only one ballot drawing function: $D_{0,1}$. Clearly $D_{0,1}(i) = hg(i, N, ceiling(\frac{N-1}{2}), n_1)$, where hg is the hypergeometric distribution.

Since no other ballots have been drawn we have $C_1 = D_{0,1}$, and the k^+ is calculated using the risk schedule.

For further rounds, we have $D_{b,j}(i) = hg(i, N, ceiling(\frac{N-1}{2}) - b, rnd(j) - rnd(j-1))$. The calculation of C_j involves $D_{b,j}$ and C_{j-1}. Observe that, since the audit proceeded to round j, the chance of having accumulated $\geq k^+_{j-1}$ ballots for the reported winner by round $j-1$ is 0.

This motivates the truncation of C_{j-1} for the purposes of computing further rounds. Define $C'_{j-1}(i) = C_{j-1}(i), \forall i \leq k^+_{j-1}$, and 0 otherwise.

The probability of having accumulated i ballots (for the reported winner) by round j (inclusive) can be represented as the sum of all probabilities of combinations of drawing b ballots before round j and t ballots during round j, where $b + t = i$. Then

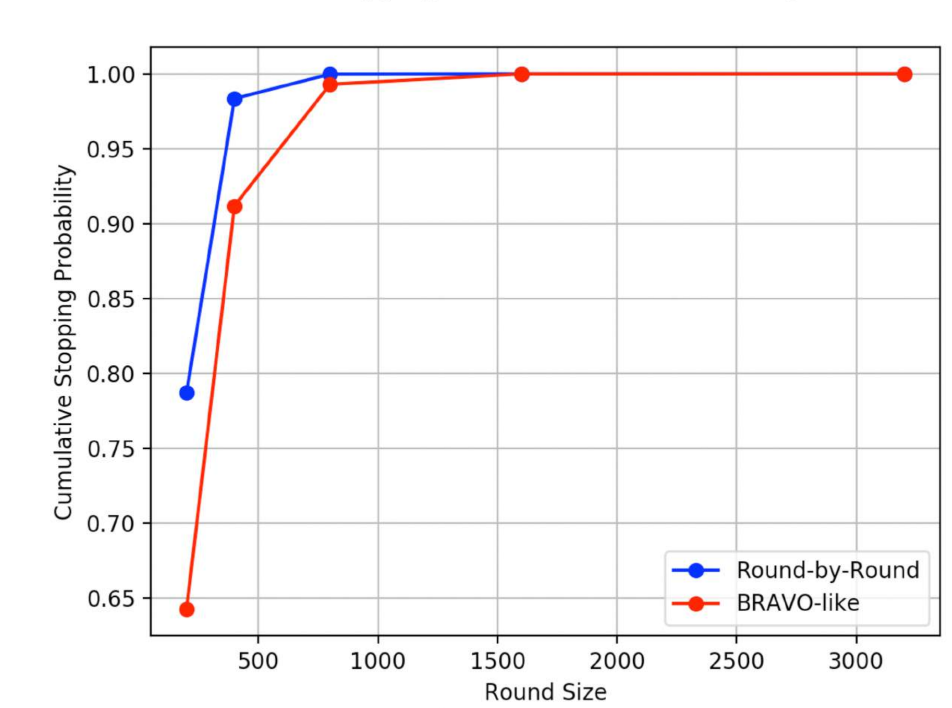
$$C_j(i) = \sum_{b+t=i} D_{b,j}(t) \times C'_{j-1}(b)$$

and k^+_j is calculated using the risk schedule.

Stopping Probabilities Comparison

The stopping probability of an *RLA* is the probability that it stops and certifies a *correctly called* election. High stopping probabilities, for a fixed risk limit, are indicative of an efficient audit.

Audit Stopping Probabilities for 20% Margin



This round-by-round audit (in blue) has a decreasing geometrically risk schedule, with common ratio 1/2.

Note: N: 100,000, Risk-limit: 5%

Handling Invalid Ballots

- During a 2-candidate election audit, we find invalid ballots: ballots not clearly marked for exactly one of the candidates.
- Traditional RLAs *do not* utilize the information gained when such ballots are found.
- The Round-by-Round RLA can incorporate these into the stopping rule calculation while remaining risk-limiting.
- In some cases, this may produce a more efficient audit.

Compute the Round-by-Round RLA with invalid ballots as follows:

1. Take $rnd(j)' = rnd(j) - r_j$ and $N' = N - r_j$
2. Perform audit calculations with $rnd(j)'$ and N'
3. Maintain N' and $rnd(j)'$ as new N and $rnd(j)$ for computation of successive rounds

Audit Stopping Rules: 20% Margin, 20% Invalid Ballots