Risk-Limiting Audits by Stratified Union-Intersection Tests of Elections (SUITE)

Kellie Ottoboni
Philip B. Stark, Mark Lindeman, and Neal McBurnett

University of California, Berkeley
Department of Statistics

Risk-limiting audits

Statistical check that tabulation errors would not change the electoral outcome.

Risk limit: chance of failing to correct a wrong outcome

https://www.verifiedvoting.org/resources/voting-equipment/

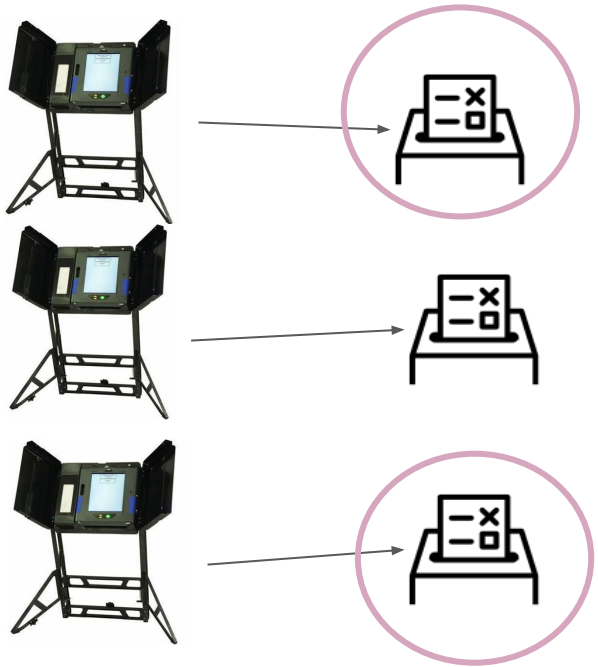
RLAs are hypothesis tests.

H_0: The reported winner is **wrong**.

Risk = P-value

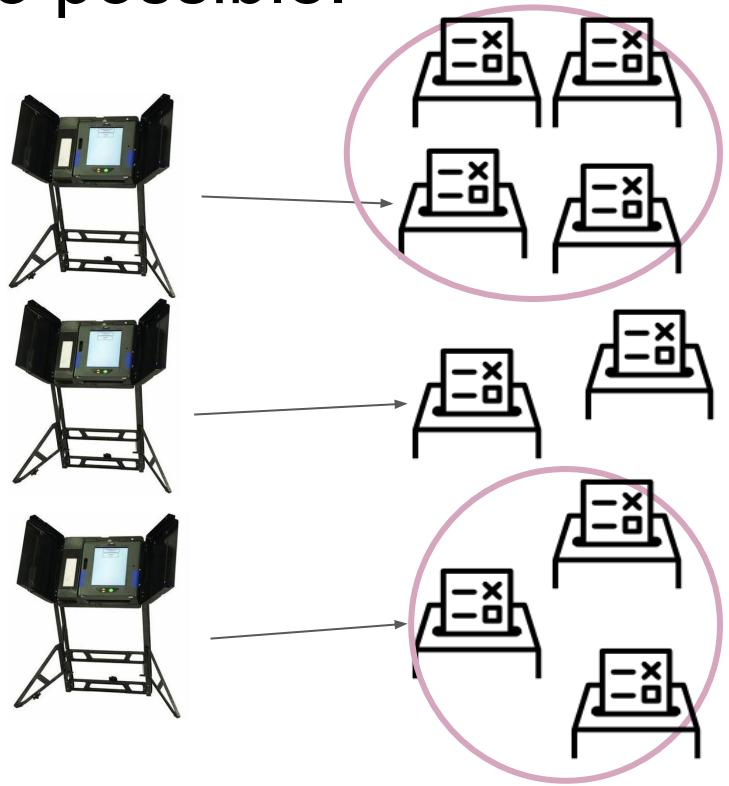
Ballot-level comparison audits

- Voting machines must export a **cast vote record (CVR)**
- Every ballot can be linked to the CVR
- Compare manual interpretation of ballots to CVR



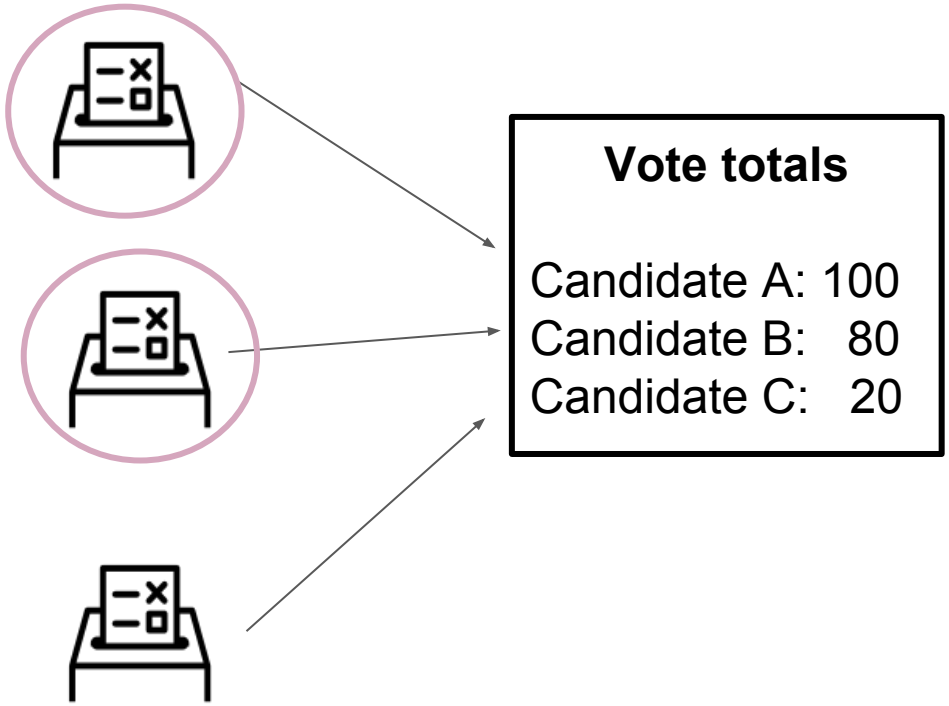
Batch-level comparisons are also possible.

- Compare machine tally with audit tally



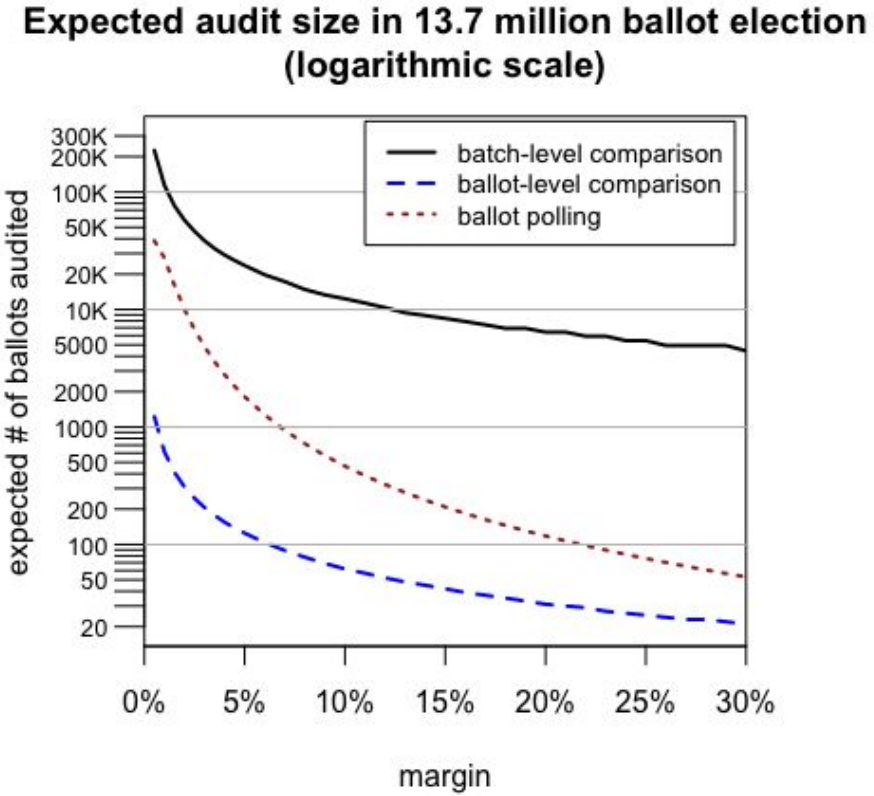
Ballot-polling audits

- Like drawing marbles from a jar and estimating proportions
- No CVR used
- Less efficient than ballot-level comparisons



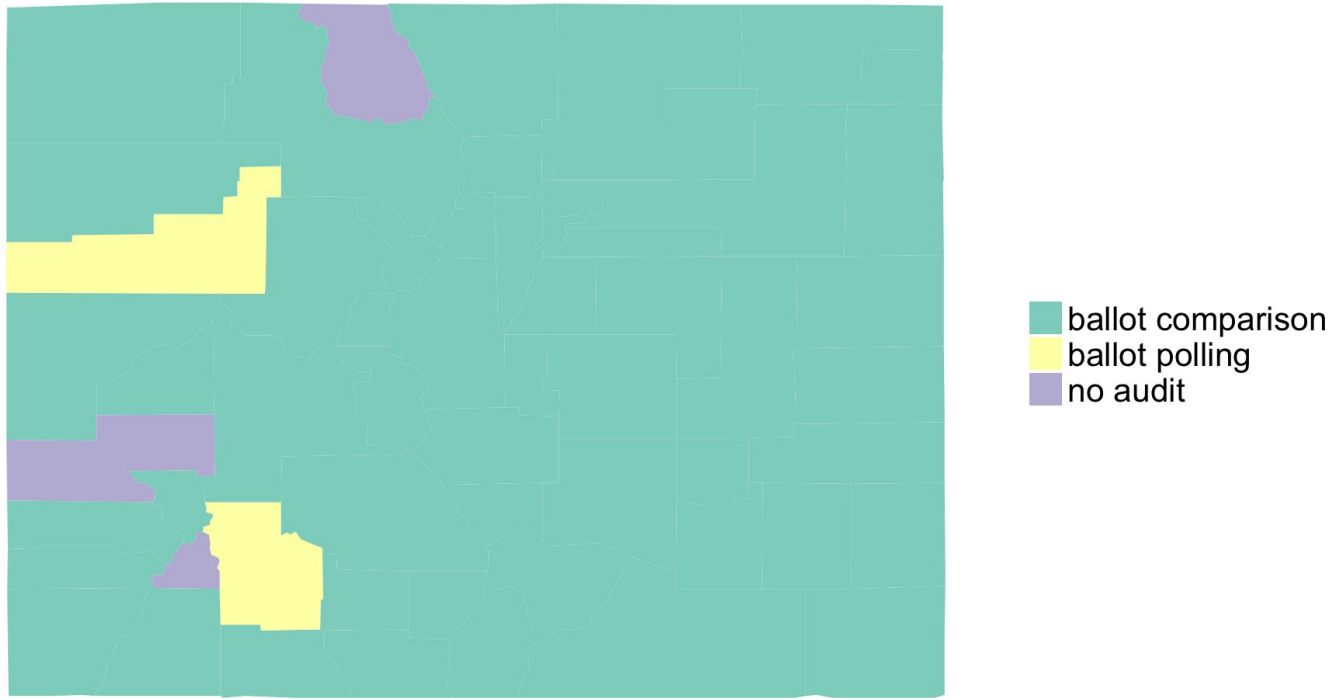
Workload depends on contest margins and method.

(These estimates assume that there are no errors!)



Colorado RLAs

Existing possibilities: ballot polling or batch-level comparisons everywhere



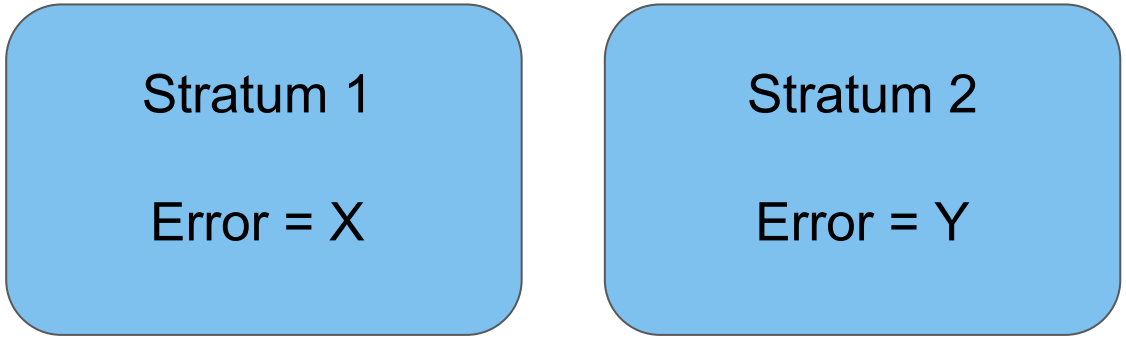
Stratified random sampling

Stratified sampling

- Partition a population into non-overlapping groups and draw independent random samples from those groups
- Useful when there are separate pools of ballots:
  - Contests that span multiple jurisdictions
  - Vote-by-mail, provisional ballots, and ballots cast in person
  - Ballots cast on heterogeneous voting equipment
- Higgins et al. (2011) considered RLAs for stratified samples using a particular auditing method and test statistic

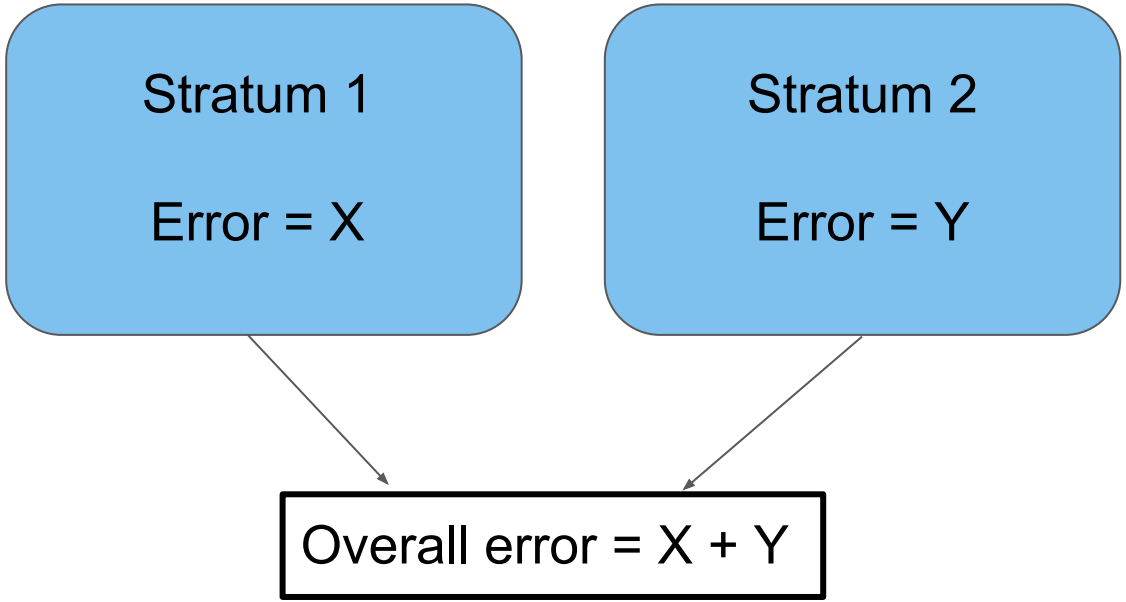
SUITE

Idea: sample strata independently, then combine the audit results.



SUITE

Idea: sample strata independently, then combine the audit results.



SUITE

Idea: sample strata independently, then combine the audit results.

$$\text{Overall error} \geq \text{Margin}$$

$$\text{if and only if}$$

$$\text{there exists } (\lambda_s)_{s=1}^{S} \text{such that}$$

$$\text{Error in stratum } s \geq \lambda_s \text{Margin}$$

SUITE

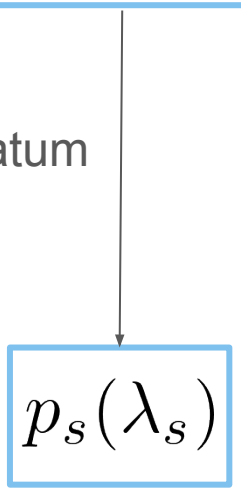
Idea: sample strata independently, then combine the audit results.

$$H_0 \iff \cup_\lambda \cap_{s=1}^{S} H_{0s}(\lambda_s)$$

Stop the audit by showing that no such λ exists.

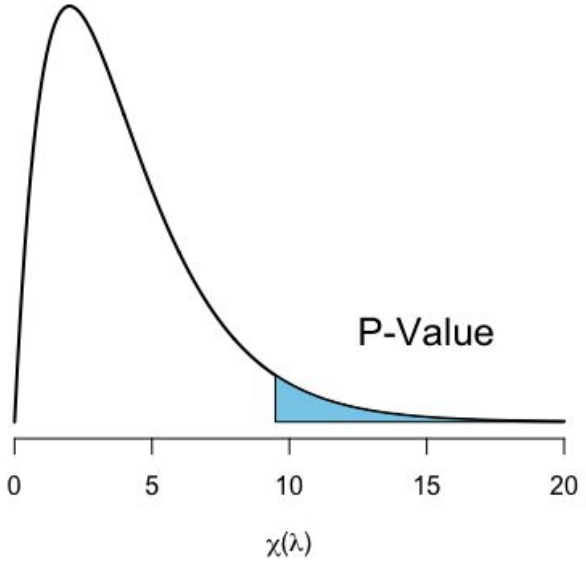
$$H_0 \iff \cup_{\lambda} \cap_{s=1}^{S} H_{0s}(\lambda_s)$$

- Method is agnostic to the auditing strategy in each stratum
- For each stratum-level hypothesis, obtain a p-value



$$H_0 \iff \cup_\lambda \boxed{\cap_{s=1}^{S} H_{0s}(\lambda_s)}$$

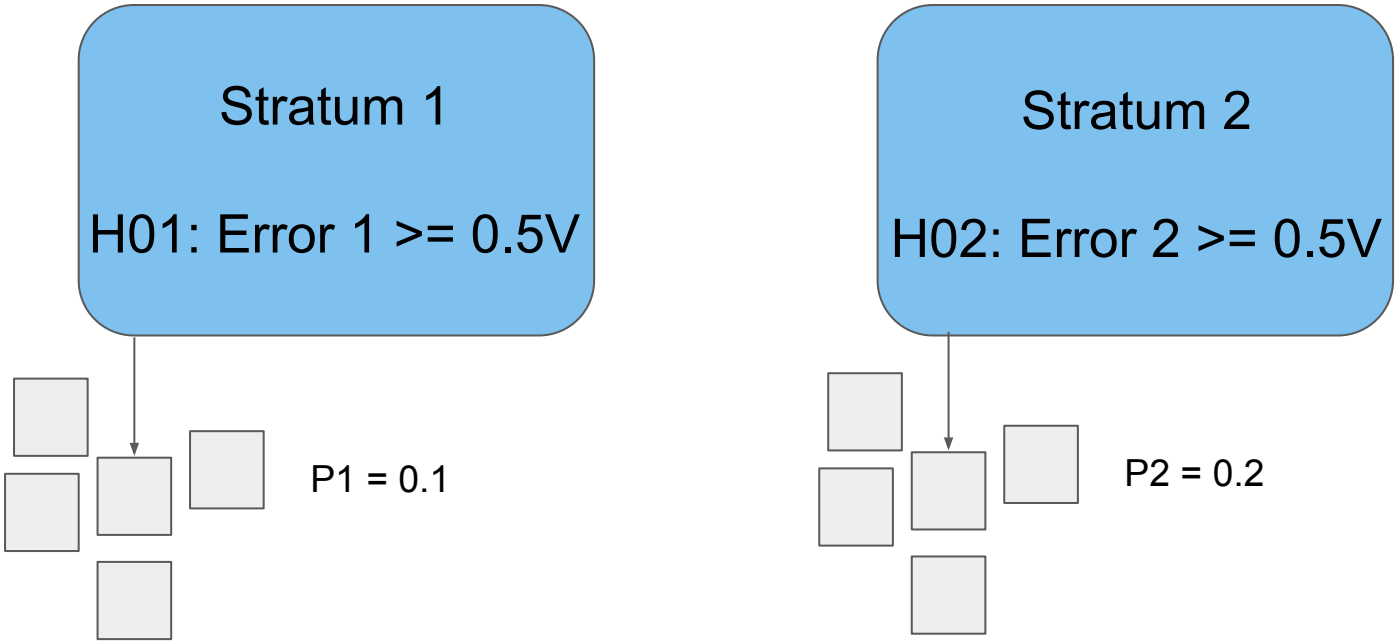
$$\chi(\lambda) = -2\sum_{s=1}^{S} \ln p_s(\lambda_s)$$



χ_{2S}

SUITE

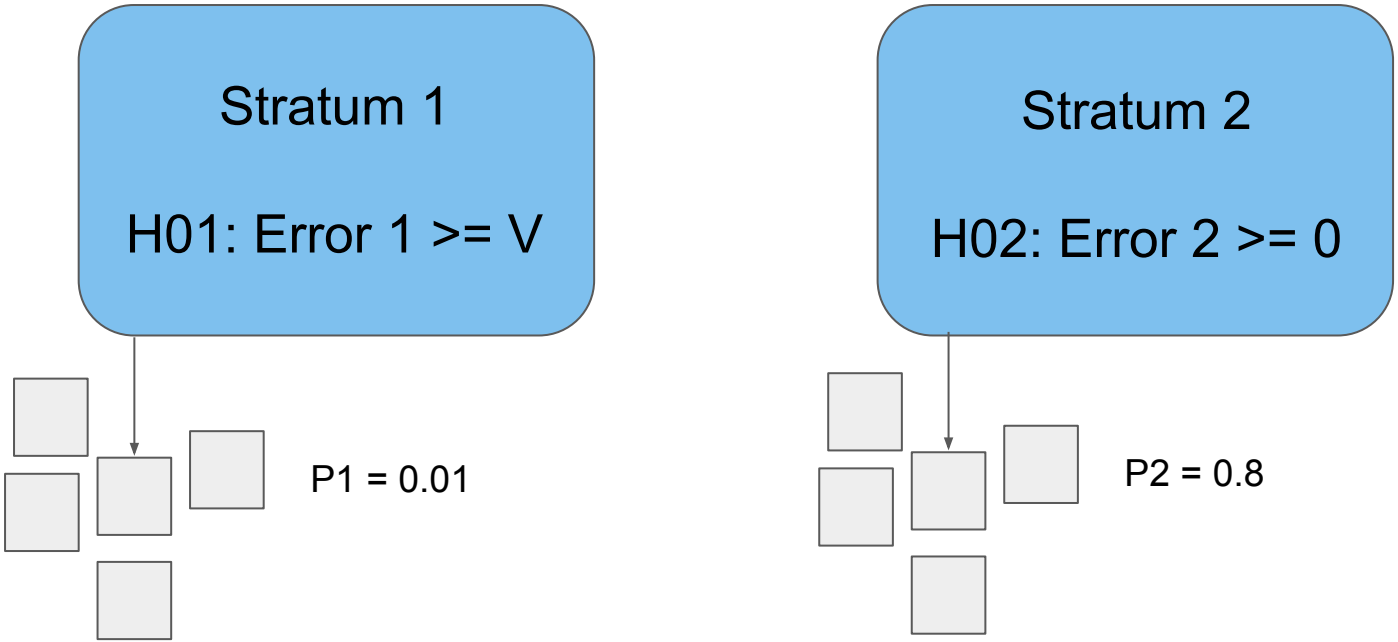
H0: Total error >= V



$$\mathbb{P}\left(\chi_4 >= -2\ln(0.1) - 2\ln(0.2)\right) = 0.09$$

SUITE

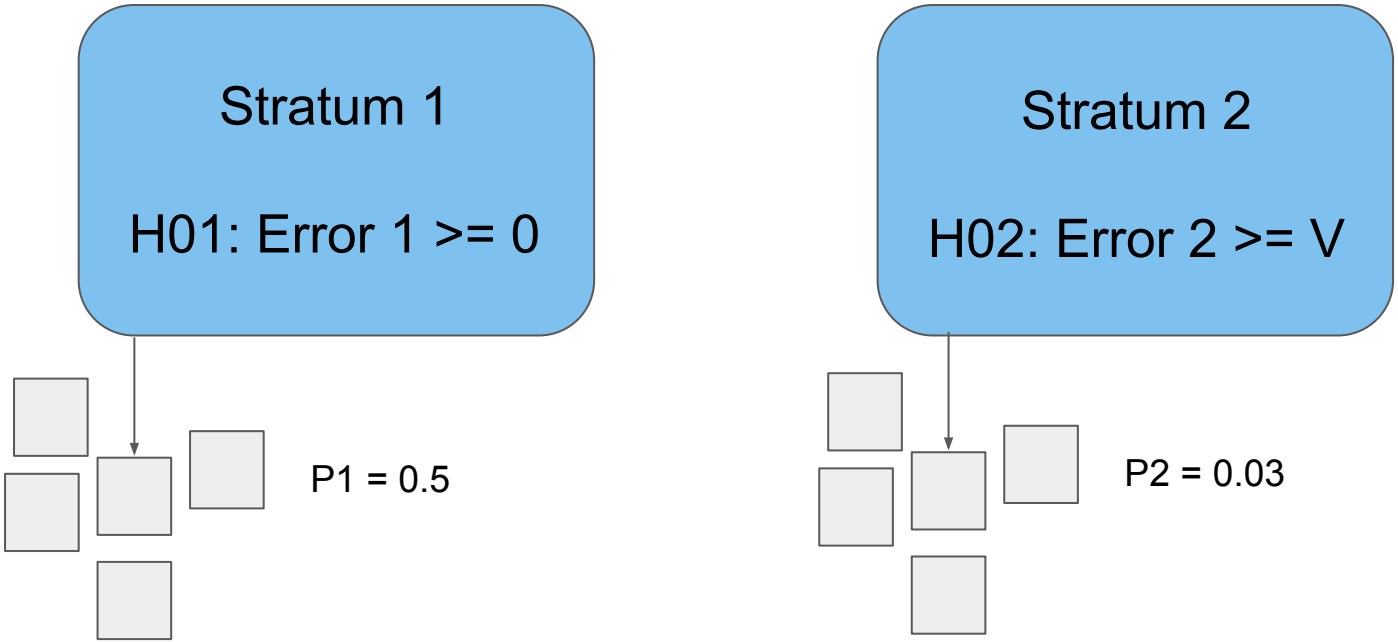
H0: Total error >= V



$$\mathbb{P}\left(\chi_4 >= -2\ln(0.01) - 2\ln(0.8)\right) = 0.05$$

SUITE

H0: Total error >= V



$$\mathbb{P}\left(\chi_4 >= -2\ln(0.5) - 2\ln(0.03)\right) = 0.08$$

$$H_0 \iff \boxed{\cup_\lambda} \cap_{s=1}^{S} H_{0s}(\lambda_s)$$

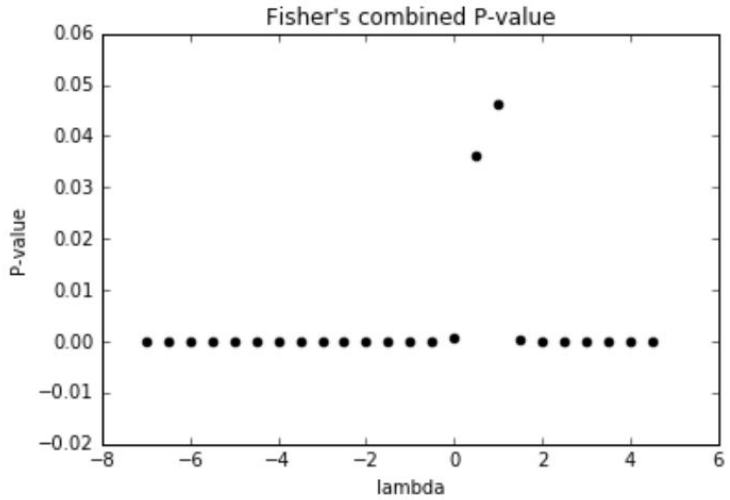
To stop the audit, we must always reject $H_{0s}(\lambda_s)$ at risk limit α.

$$H_0 \iff \boxed{\cup_\lambda} \cap_{s=1}^{S} H_{0s}(\lambda_s)$$

To stop the audit, we must always reject $H_{0s}(\lambda_s)$ at risk limit α.

Maximize Fisher's combined P-value:

- Approximate via grid search of λ values

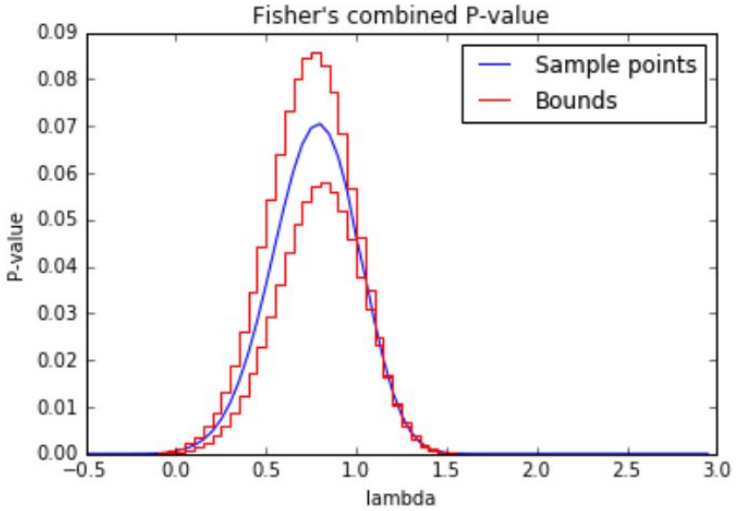


$$H_0 \iff \boxed{\cup_\lambda} \cap_{s=1}^{S} H_{0s}(\lambda_s)$$

To stop the audit, we must always reject $H_{0s}(\lambda_s)$ at risk limit α.

Maximize Fisher's combined P-value:

- Approximate via grid search of λ values
- For S=2, construct piecewise constant upper and lower bounds for the P-value



SUITE

1. Divide ballots into non-overlapping strata.

2. Draw initial samples from each stratum.

3. Maximize Fisher's combined P-value over all error partitions λ.

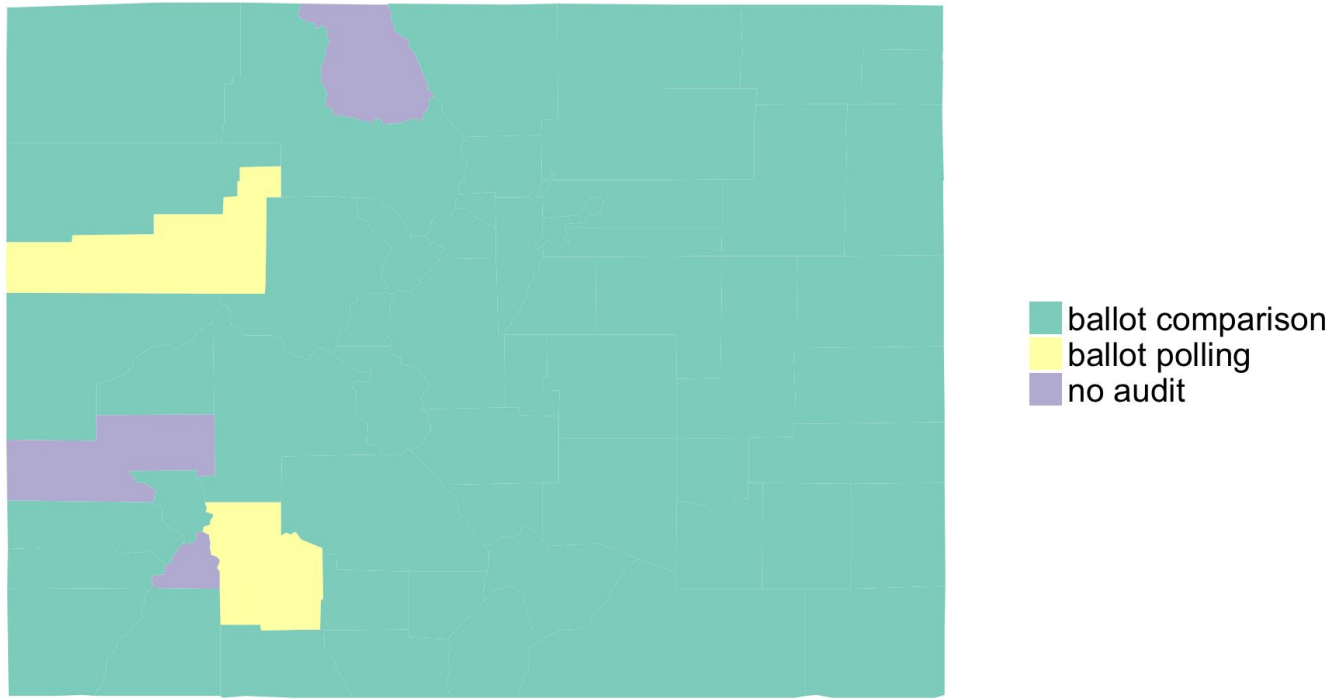
4. If the maximum is above the risk limit, escalate the audit and repeat.

SUITE applied:

Colorado Case Study

Colorado RLAs

Existing possibilities: ballot polling or batch-level comparisons everywhere



Colorado RLAs

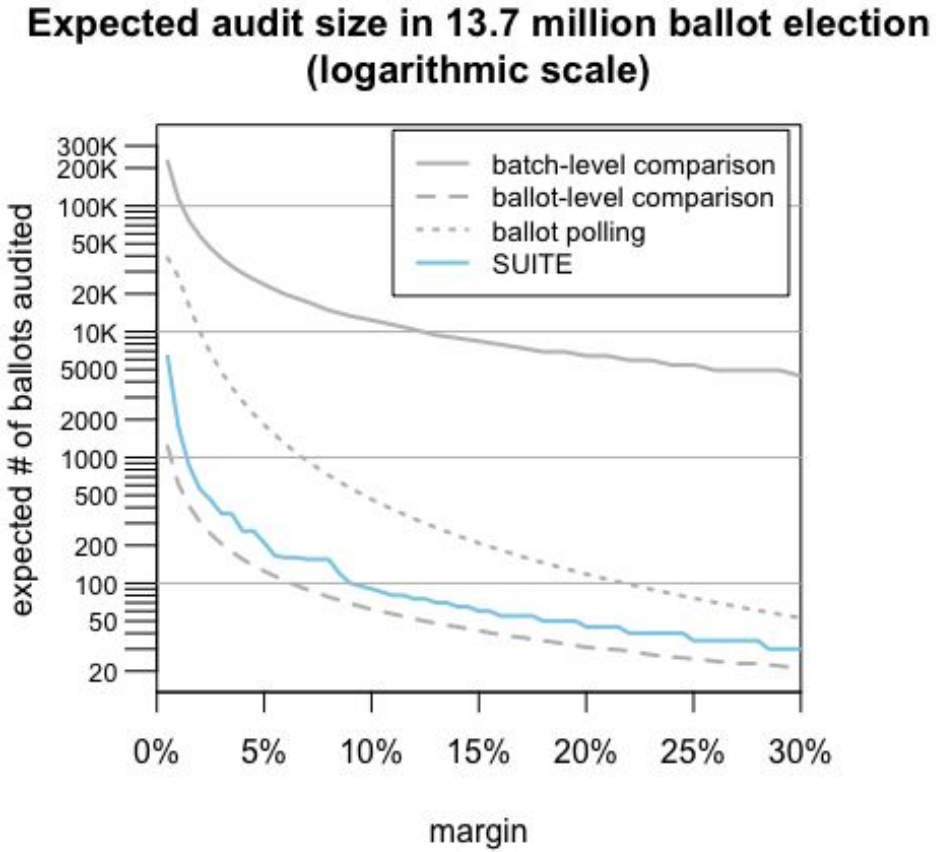
Existing possibilities: ballot polling or batch-level comparisons everywhere

Solution using SUITE:

- **Two strata**: one with ballot polling counties, one with the rest
- Requires modifying existing tests for ballot polling and ballot comparison

SUITE in Colorado

(These estimates assume that there are no errors!)



- SUITE is a general risk-limiting audit for stratified samples
- Agnostic to audit strategy in each stratum
- Immediate application for statewide contests in the U.S.

Thanks!


kellieotto@berkeley.edu


www.kellieottoboni.com


@kellieotto