Simulating a Ballot-Polling Audit with Cards and Dice

Multidisciplinary Conference on Election Auditing

MIT, 7-8 December 2018

P.B. Stark

University of California, Berkeley stark@stat.berkeley.edu

Version: December 7, 2018

Risk-limiting audits (RLAs) offer a statistical guarantee: if a full manual tally of the paper ballots would show that the reported election outcome is wrong, an RLA has a known minimum chance of leading to a full manual tally. The *risk limit* is the maximum chance the audit won't lead to a full hand tally, if the reported outcome is wrong.

Simulating a ballot-polling RLA

Ballot polling is an approach to auditing that involves selecting individual ballots at random from the cast ballots and interpreting the votes on the ballots manually. Ballot polling does not involve comparing the human interpretation to anything else. (In contrast, *comparison audits* compare the human interpretation to the voting system's interpretation.)

If the sample shows substantially more votes for the reported winner than for any reported loser, that is evidence that the reported winner really won. Putting a number on "substantially" to enforce the risk limit requires some deep math, but the resulting arithmetic is very simple: just multiplication, addition, and subtraction.

In this exercise, we will simulate a ballot-polling audit with 10% risk limit. That is, if the reported outcome is wrong, there is at least a 90% chance that the audit will lead to a full hand count, and at most a 10% chance that it won't.

Set up a correct election outcome

Your group should have two decks of cards.

1. Remove the jokers and the hearts from both decks and the diamonds from one deck. This leaves 65 cards, of which 52 are black and 13 are red. Each card represents a ballot.
2. Arrange the 65 cards, face down, in 6 "precincts," labeled "A" through "F," each containing at least 5 ballots.[1] Note how many ballots are in each precinct, and make a "ballot manifest" on a sheet of paper. Your manifest should look something like this:

   ```
   A: 7
   B: 15
   C: 12
   D: 9
   E: 12
   F: 10
   ```

3. Do "ballot accounting" to ensure that the numbers in your ballot manifest add up to 65.
4. Make a "ballot look-up table" to help locate randomly selected ballots by finding the cumulative number of ballots across batches. For the example manifest, this would be:

   ```
   A: 7    1-7
   B: 15   8-22
   C: 12   23-34
   D: 9    35-43
   E: 12   44-55
   ```

[1] In the U.S., precinct and batch sizes vary from a handful of votes to roughly 2,000.

```
F: 10   56-65
```

5. In this audit simulation, the reported results will be correct. Color will represent a candidate. The true contest results are 52 black to 13 red, a margin of (52-13)/65 = 39/65 = 60%. Black is the reported winner. The audit checks whether the reported winner is correct (the *electoral outcome*), not whether the vote tally is exactly correct. That is, the audit stops if and when it finds sufficiently strong evidence that there really are more black cards than red cards and that continuing to count would be a waste of time.

6. The rule for when the audit can stop is as follows:

   i. Set the running total to 0.
   ii. Generate a random number between 1 and 65. Record the number.
   iii. Retrieve the corresponding card, using the ballot look-up table. Record the color.
      - If the card is black, add 5
      - If the card is red, subtract 9.75 (if you prefer to subtract 10, that's still safe)
   iv. If the running total is greater than 24.5, stop. Otherwise, put the card back in its original place in the "precint" and return to step (ii), or optionally do a full hand count.

   - You might end up selecting the same ballot more than once as the audit progresses. You only have to inspect it once, but you should update the total every time it is selected. For instance, if you select the same black card three times, it contributes 15 to the sum, not 5.)

7. Why does this rule make sense?

   - Seeing a black card adds to the evidence that black won. Hence, we add to the running total, bringing it closer to the threshold for stopping the audit.
   - Conversely, seeing a red card reduces the strength of the evidence that black won. Hence, we subtract from the running total, delaying the end of the audit.
   - A red card casts more doubt on the outcome than a black card removes: you have to see about two black cards to compensate for each red card.
   - If first 5 cards selected are all black, the total will be 25, and the audit can stop. It would be very unlikely for the first 5 cards to be black if the true outcome were a tie, or if red had won.
   - If one of the first 5 cards drawn is red, the audit will have to inspect at least 7 cards before it can stop.
   - On average, the audit will have to examine 14 cards before it can stop.

8. Choose a random sampling method. Either you can roll dice for every selection, or roll dice to generate a "seed" and use the pseudo-random number generator: https://www.stat.berkeley.edu/~stark/Java/Html/sha256Rand.htm Regardless, you will be generating random numbers between 1 and 65, and inspecting the corresponding ballot. For instance, for the ballot manifest above, if the number is 27, you would inspect the 5th ballot in precinct C.

   - If you want to roll dice for every selection, use two dice of different colors. Designate one color to be the tens digit and one to be the units digit. To select a ballot, roll both dice. If the number is between 1 and 65, inspect the corresponding ballot. If the number is 0 or greater than 65, roll again. (You will end up discarding about 1/3 of the rolls.)
   - If you want to use the pseudo-random number generator, roll 10-sided dice 20 times to generate a 20-digit "seed," record the seed in your "audit log," and type the seed into the appropriate box on the webpage. Type "65" into the box labeled "Number of objects from which to sample." Click the "reset" button. Each time you click the "draw sample" button, the software will generate another number between 1 and 65.

9. Audit! The number of ballots you will have to inspect to confirm the outcome is random. It could be as few as 5. On average, you will need to inspect 14 ballots to confirm the contest results.

Set up the second election and audit

1. Take the 37 of the 39 hearts and diamonds you set aside and arrange them into 3-5 additional "precincts." Remove two of the black cards from the existing precincts. This should leave 50 red cards and 50 black cards.

Update your ballot manifest and look-up table to show all 100 cards. Double-check that the manifest lists 100 cards in all. Imagine that the reported outcome was that black won, with 60% of the vote, as before; but in reality, this election is a tie. A risk-limiting audit at 10% risk limit will have at most a 10% chance of stopping without a full hand count.

2. Start the audit over: set the running total to zero.

3. Using either dice or the online pseudo-random number generator, select numbers between 1 and 100, record them, find the corresponding ballot using the look-up table, inspect it, record the color, and update the running total according to the color of the card (add 5 if the card is black; subtract 9.75 or 10 if the card is red). To generate numbers between 1 and 100 using dice, use dice of two colors. Designate one of them to be the tens place and the other to be the ones place. Roll both dice. That gives a number between 0 and 99. Add 1 to the result to get a random number between 1 and 100.

4. Continue to select cards and update the running total until you convince yourself that the sum is unlikely ever to get to 24.5 (generally, the running total will tend to get smaller, not bigger). When you get bored, "cut to the chase" and do a full hand count to figure out the correct election outcome. Because the risk limit is 10%, there is a 10% chance that the audit will stop before there is a full hand count, confirming an incorrect outcome.

- If we wanted the risk limit to be 5% instead of 10%, everything would stay the same except the threshold, which would increase from 24.5 to 31.9: we would have to see (on average) 1.5 additional black ballots before the audit could stop.