Justice and Inequalities

Department of Statistics and Operations Research
Tel Aviv University
Tel Aviv, Israel
13 April 2010

Philip B. Stark
http://statistics.berkeley.edu/~stark

This document:
http://statistics.berkeley.edu/~stark/Seminars/tau10.pdf

Abstract

I will discuss some inference problems in election auditing and litigation that can be solved using probability inequalities. The lead example, illustrated with case studies in auditing elections and estimating damages in civil litigation, is to construct nonparametric one-sided confidence bounds for the mean of a nonnegative population. If time permits, I will also discuss a contested election in which a simple probability inequality provided evidence the court found persuasive. This seminar is partly a plea for help from probabilists: I hope someone in the audience can point me to inequalities that are sharper than those I'm using.

Election Auditing: Any way of counting votes makes mistakes.

If there are enough mistakes, apparent winner could be wrong.

If there's an audit trail that reflects the right outcome, can ensure big chance of fixing wrong outcomes.

Crucial question: when to stop counting.

Solution: If there's compelling evidence that outcome is right, stop; else, audit more.

Risk-Limiting Audits

> If the electoral outcome is wrong, there's a known minimum chance of a full hand count (which fixes it), *no matter what caused the outcome to be wrong*.

The *risk* is the largest chance that an outcome that is wrong won't be fixed.

"Wrong" means the outcome isn't what a full hand count would show.

More delicate in parliamentary elections than in "first past the post" elections: Did each party get the right number of seats?

Role of statistics: Less counting when the outcome is right, but still a big chance of a full hand count when outcome is wrong.

Essential that voters create a durable audit trail that reflects the true outcome.

Essential that voting systems enable auditors to access reported results (total ballots, counts for each candidate) in auditable batches.

Essential to perform "ballot accounting" to ensure that no ballots appeared or disappeared.

Essential to select batches at random, *after* the results are posted or "committed." (Can supplement with "targeted" samples.)

Need a plan for when to count more ballots, possibly leading to full hand count. "Explaining" or "resolving" isn't enough. Plan must ensure that the chance of a full hand count is high whenever the outcome is wrong.

Compliance audits vs. materiality audits.

Quantifying the Evidence the Audit Sample Gives

What is the biggest chance that—if the outcome is wrong—the audit would have found “as little” error as it did?

Chance can be big *even if no errors are found*—if the sample is small or the margin is small.

Don’t stop counting until that chance is small!

Sequential test of null hypothesis that the outcome is *wrong* at significance level α controls the risk to α.

Notation

1 contest at a time (can combine C contests), vote for $\leq f$ of K candidates.

Results subtotaled in N auditable batches, $p = 1, \ldots, N$.

a_{kp}: actual votes for candidate k in batch p.

v_{kp}: reported votes for candidate k in batch p.

$A_k \equiv \sum_{p=1}^{N} a_{kp}$; $V_k \equiv \sum_{p=1}^{N} v_{kp}$.

$V_{w\ell} \equiv V_w - V_\ell$: margin of w over ℓ

$\mathcal{W}$: the f apparent winners; $\mathcal{L}$: the $K - f$ apparent losers.

If $w \in \mathcal{W}$ and $\ell \in \mathcal{L}$, then $V_{w\ell} > 0$.

Summarizing error

Candidate w really beat candidate ℓ if $A_w - A_\ell > 0$.

Relative overstatement of the margin between $w \in \mathcal{W}$ and $\ell \in \mathcal{L}$ in batch p:

$$e_{w\ell p} = \frac{v_{wp} - v_{\ell p} - (a_{wp} - a_{\ell p})}{V_{w\ell}}.$$

The outcome of the race is correct if $\forall w \in \mathcal{W}$, $\ell \in \mathcal{L}$,

$$\sum_{p=1}^{N} e_{w\ell p} < 1.$$

Apparent winner(s) are real winner(s) if no margin was overstated by 100% or more of *that* margin.

Maximum relative overstatement of pairwise margins

Let

$$e_p \equiv \max_{w \in \mathcal{W}, \ell \in \mathcal{L}} e_{w\ell p}.$$

To audit C contests simultaneously,

$$e_p \equiv \max_c \max_{w \in \mathcal{W}_c, \ell \in \mathcal{L}_c} e_{w\ell p}.$$

Now

$$\max_{w \in \mathcal{W}, \ell \in \mathcal{L}} \sum_p e_{w\ell p} \le \sum_p \max_{w \in \mathcal{W}, \ell \in \mathcal{L}} e_{w\ell p} = \sum_p e_p.$$

All apparent outcomes are right if

$$E \equiv \sum_{p=1}^{N} e_p < 1.$$

Testing $E \geq 1$

How strong is the evidence that $E < 1$? Maximum P-value.

Need upper bounds on $\{e_p\}$; otherwise P-value for $E \geq 1$ large unless we audit most batches.

b_p: bound on valid votes for any candidate in batch p.*

Reported margin vs. all b_p votes really for ℓ:

$$
\begin{aligned}
e_{w\ell p} &\leq \frac{v_{wp} - v_{\ell p} - (0 - b_p)}{V_{w\ell}} \\
&= \frac{b_p + v_{wp} - v_{\ell p}}{V_{w\ell}}.
\end{aligned}
$$

*E.g., from voter registrations, accounting of ballots, pollbook signatures.

A priori batch error bounds

Define

$$u_p \equiv \max_c \max_{w \in \mathcal{W}_c, \ell \in \mathcal{L}_c} \frac{b_p + v_{wp} - v_{\ell p}}{V_{w\ell}}.$$

Then

$$e_p \leq u_p.$$

Surprisingly controversial among EI advocates.
Less controversial among elections officials. Hmmm.

Extends to simultaneous audits of several races (MARROP); controls FWER with cost comparable to controlling PCER.

Sampling Designs

Simple

Stratified (by county, voting method, other)

PPEB

NEGEXP

Stratified PPEB?

Sampling scheme affects choice of test statistic—analytic tractability

Weighted max for simple & stratified sampling.

More efficient choices possible for PPEB.

PPEB sampling

$$\text{total error bound: } U \equiv \sum_{p=1}^{N} u_p.$$

$$\text{“taint” of batch } p\text{:} \quad t_p \equiv e_p/u_p \leq 1.$$

Draw n times iid, chance u_p/U of drawing batch p.

$T_j = t_p$ if batch p is selected in draw j

$$\mathbb{E}T_j \equiv \sum_{p=1}^{N} t_p \frac{u_p}{U} = \frac{E}{U}.$$

Outcome must be right if $\mathbb{E}T_j < 1/U$.

One-sided test or confidence bound for expected value of RV bounded on one side

Test whether mean of upper-bounded r.v. $< 1/U$ from iid sample.

Canonical form:
$\{X_j\}$ iid, $\mathbb{P}\{X \ge 0\} = 1$.
Find lower confidence bound for $\mathbb{E}X_1$

Transform via $X_j \equiv 1 - T_j$.

Binning

Define new RVs:

$$\tilde{X}_j \equiv \begin{cases} 0, & X_j < t \\ t, & X_j \geq t \end{cases}, \quad j = 1, \ldots, n.$$

Let $p \equiv \mathbb{P}\{\tilde{X}_1 > 0\}$.

Then

$$\mathbb{E}\tilde{X}_1 = pt \leq \mathbb{E}X_1.$$

$$\sum_{j=1}^{n} \tilde{X}_j \sim t \times \mathrm{Bin}(n, p).$$

Confidence bound on $\mathbb{E}X_1$ from confidence bound on p.

Tests from Markov's Inequality

Markov's inequality: If $\mathbb{P}\{X \geq 0\} = 1$, then

$$\mathbb{P}\{X \geq \mathbb{E}X/\tau\} \leq \tau.$$

Implies that if $\{X_j\}_{j=1}^n$ iid, $\mathbb{P}\{X_1 \geq 0\} = 1$,

$$\mathbb{P}\left\{\cap_j \{X_j \geq \mathbb{E}X_1/\tau\}\right\} \leq \tau^n.$$

(i) Reject the hypothesis $\mathbb{E}X_1 \leq \mu$ at significance level α on observing $\bar{X} \equiv \frac{1}{n}\sum_{j=1}^n X_j = x$ if

$$x \geq \mu/\alpha.$$

(ii) Reject the hypothesis $\mathbb{E}X_1 \leq \mu$ at significance level α on observing $X^- \equiv \min_{j=1}^n X_j = x$ if

$$x \geq \mu/\alpha^{1/n}.$$

MDKW confidence bounds

$\mathbb{P}(X_k \geq 0) = 1$; $\{X_k\}_{k=1}^n$ iid, cdf F.

$$\widehat{F}_n(x) \equiv \frac{1}{n} \sum_{k=1}^{n} 1_{x \geq X_k}$$

$$\mathbb{E}X \equiv \int x dF(x)$$

One-sided Kolmogorov-Smirnov statistic

$$D_n = \sup_x (F(x) - \widehat{F}_n(x)).$$

MDKW inequality:

$\mathbb{P}_F\{D_n > \chi\} \le \exp(-2n\chi^2)$ for $\exp(-2n\chi^2) \le \frac{1}{2}$.

Hence $\mathbb{P}_F\left\{D_n > \sqrt{-\frac{\ln \alpha}{2n}}\right\} \le \alpha$.

Shift as much mass as possible to the left: Define

$$\tilde{F}_{n,\alpha}(x) \equiv 1 \wedge \left(\widehat{F}_n(x) + \sqrt{-\frac{\ln \alpha}{2n}}\right).$$

Then $\mathbb{P}_F\left\{\int_0^\infty x d\tilde{F}_{n,\alpha}(x) \le \mathbb{E}X\right\} \ge 1 - \alpha$.

$$\int_0^\infty x d\tilde{F}_{n,\alpha}(x)$$

is a $1-\alpha$ lower confidence bound for $\mathbb{E}X$. Easy to compute.

Martingale: $X_1, X_2, \ldots$ such that $\mathbb{E}|X_j| < \infty$ and

$$\mathbb{E}(X_{j+1}|X_1, \ldots, X_j) = X_j \text{ (a.s.)}.$$

If $X_1, X_2, \ldots$ is a Martingale and $x > 0$,

$$\mathbb{P}\left(\max_{1\le j\le k} X_j > x\right) \le \mathbb{E}|X_k|/x.$$

IID rvs to martingales

Suppose $\{Y_j\}$ iid, $\mathbb{P}\{Y_j \ge 0\} = 1$; $\mathbb{E}Y_j = \mathbb{E}|Y_j| = \mu < \infty$.

Let

$$X_j \equiv \prod_{i \le j} Y_i / \mu.$$

Note $X_{j+1} = X_j \cdot Y_{j+1}/\mu$, $\mathbb{E}X_j = 1$, and

$$\mathbb{E}(X_{j+1}|X_1, \ldots, X_j) = X_j \mathbb{E}Y_{j+1}/\mu = X_j.$$

So, $X_1, X_2, \ldots$ is a Martingale.

Kaplan-Markov Martingale P-value & Bound

Substitute definition of X_j and set $x = 1/P$:

$$\mathbb{P}\left(\max_{1\le j\le k}\prod_{i=1}^{j} Y_j/\mu > 1/P\right) \le P$$

Lower $1-\alpha$ confidence bounds for $\mathbb{E}X$ from $\{X_k\}$.
1-permutation bound:

$$\max_{\ell=1,\ldots,n}\left(\alpha\prod_{i=1}^{\ell} X_i\right)^{1/\ell}$$

all-permutation bound:

$$\max_{\ell=1,\ldots,n}\left(\frac{\alpha\prod_{k=1}^{\ell} X_{(n-k+1)}}{\binom{n}{\ell}}\right)^{1/\ell},$$

where $X_{(k)}$ is the kth order statistic of $\{X_j\}$.

Method	Observed Taints					
	clean	0.01	0.01 0.01	0.02	0.01 0.03	-0.05 $\times$5 0.05 $\times$5
$n = 10$, $U = 5$						
Bin, $t = 0.01$	0.119	0.119	0.119	0.401	0.401	0.995
Bin, $t = 0.02$	0.131	0.131	0.131	0.131	0.427	0.996
Markov-max	0.107	0.119	0.119	0.131	0.146	0.179
MDKW	0.449	0.453	0.457	0.457	0.464	0.484
KM	0.107	0.108	0.110	0.110	0.112	0.109
$n = 20$, $U = 5$						
Bin, $t = 0.01$	0.014	0.014	0.014	0.081	0.081	0.830
Bin, $t = 0.02$	0.017	0.017	0.017	0.017	0.095	0.854
Markov-max	0.012	0.014	0.014	0.017	0.021	0.032
MDKW	0.202	0.204	0.205	0.205	0.208	0.234
KM	0.012	0.012	0.012	0.012	0.012	0.012
$n = 30$, $U = 10$						
Bin, $t = 0.01$	0.057	0.057	0.057	0.229	0.229	0.950
Bin, $t = 0.02$	0.078	0.078	0.078	0.078	0.285	0.968
Markov-max	0.042	0.057	0.057	0.078	0.106	0.198
KM	0.042	0.043	0.043	0.043	0.044	0.043
$n = 40$, $U = 15$						
Bin, $t = 0.01$	0.095	0.095	0.095	0.324	0.324	0.975
Bin, $t = 0.02$	0.142	0.142	0.142	0.142	0.426	0.989
Markov-max	0.063	0.095	0.095	0.142	0.214	0.493
KM	0.063	0.064	0.065	0.065	0.066	0.064

Sequential risk-limiting test

0. Calculate error bounds $\{u_p\}$, U. Set $n = 1$. Pick $\alpha \in (0, 1)$ and $M > 0$.

1. Draw a batch using PPEB. Audit it if not audited previously.

2. Find $T_n \equiv t_p \equiv e_p/u_p$, taint of the batch p just drawn.

3. Compute

$$P_n \equiv \prod_{j=1}^{n} \frac{1 - 1/U}{1 - T_j}.$$

4. If $P_n < \alpha$, stop; report apparent outcome. If $n = M$, audit remaining batches. If all batches have been audited, stop; report known outcome. Else, $n \leftarrow n + 1$ and go to 1.

This sequential procedure is risk-limiting

If outcome is wrong,

$$\mathbb{P}\{\text{stop without auditing every batch}\} < \alpha.$$

Chance $\geq 1 - \alpha$ of fixing wrong outcome by full hand count.

Remarkably efficient (in simulations).

Pilot Audits in California

Marin County 2/08 (first ever); 11/08

Santa Cruz County 11/08

Yolo County 11/08, 11/09 (2, incl. 1st single-ballot audit)

Measures requiring super-majority, simple measures, multi-candidate contests, vote-for-n contests.

Contests ranged from about 200 ballots to 121,000 ballots.

Counting burden ranged from 32 ballots to 7,000 ballots.

Cost per audited ballot ranged from nil to about $0.55.

Yolo County Measure P, November 2009

Reg. voters	ballots	precincts	batches	yes	no
38,247	12,675	31	62	3,201	9,465

VBM and in-person ballots were tabulated separately (62 batches).

For risk-limit 10%, initial sample size 6 batches; gave 4 distinct batches, 1,437 ballots.

Single-ballot auditing would save *lots* of work

Can determine the initial sample size for a Kaplan-Markov single-ballot audit even though the cast vote records (CVRs) were not available.

For risk-limit 10% would need to look at CVRs for 6 ballots.

For risk-limit 1%, 12 ballots.

Cf., 1,437 ballots for actual batch sizes.

Estimating Damages in a Labor Dispute

Class-action suit over involuntarily missed rest breaks and meal breaks. Drivers entitled to a rest break every 4 hours, meal break every 5 hours.

If break is taken, time is unpaid. If missed voluntarily, get overtime payment.

If missed involuntarily, damages based on number of involuntarily missed breaks of each kind.

Transform to problem of estimating mean of bounded r.v.

For each class member j, upper bound u_j on total hours worked, from driver logs, time sheets, payroll records, pension records.

$m_j \equiv \#(\text{meal break entitlements for member } j) \leq u_j/5$

$r_j \equiv \#(\text{rest break entitlements for member } j) \leq u_j/4$.

$U \equiv \sum_{j=1}^N u_j$: upper bound on the total hours worked by all members.

$m \equiv \sum_{j=1}^N m_j \leq U/5$.
$r \equiv \sum_{j=1}^N r_j \leq U/4$.

Probability-proportional-to-size sampling

$\mu_j \equiv m_j/(u_j/5)$ and $\rho_j \equiv r_j/(u_j/4)$. Then

$$\mu_j \in [0,1] \text{ and } \rho_j \in [0,1].$$

Draw class members independently with replacement, probability of drawing member j is $\pi_j \equiv u_j/U$ in every draw.

Members who worked more hours more likely to be sampled: Effort where the money is.

Expected distinct members in n draws:

$$\sum_{j=1}^{N} \left(1 - \left(1 - \frac{u_j}{U}\right)^n\right).$$

The more the bounds $\{u_j\}$ are skewed to the right, the smaller the expected number of distinct class members in the sample.

Taint

Let M_k and R_k be the values of μ_j and ρ_j, respectively, for the class member drawn on the kth draw, $k = 1, \ldots, n$. Now $\{M_k\}_{k=1}^n$ are iid random variables on $[0, 1]$, as are $\{R_k\}_{k=1}^n$. However, M_k and R_k are dependent and not identically distributed.

Note that

$$\mathbb{E} M_k = \sum_j \mu_j \pi_j = \sum_j \mu_j \frac{u_j}{U} = \sum_j \frac{m_j}{u_j/5} \frac{u_j}{U} = \frac{5}{U} m.$$

$(U/5)M_k$ is unbiased for m.

$(U/4)R_k$ is unbiased for r.

What does the judge need to do?

Determine sample size (ruled $n = 45$).

Determine confidence level to use, select method.

Determine, for each class member in the sample, the *actual number* of break entitlements missed involuntarily, not the fraction of break entitlements missed involuntarily. The bound on the number of entitlements for each class member was be determined from payroll and other records before the sample was drawn.

$n = 45$ draws gave 40 distinct class members.

Depositions underway.

Estimation and confidence intervals

Burden of proof is on plaintiffs: primary goal is to find lower confidence bounds for m and r, rather than unbiased estimators of m and r.

Nonparametric lower confidence bounds for the expected value of nonnegative random variables from iid samples can be applied to $\{M_k\}$ and $\{R_k\}$.

Scale those confidence bounds by $U/5$ and $U/4$ to get lower confidence bounds for m and r.

Same problem as in election auditing.

p	λ	ν	Sample Mean	n	Discount from Sample Mean: MDKW	Kaplan: 1-perm	Kaplan: all-perm
0.1	0.05	0.9	0.815	30	27.4%	21.6%	26.9%
				50	21.2%	17.6%	23.2%
				100	15.0%	14.4%	20.3%
				200	10.6%	12.8%	18.8%
0.5	0.05	0.9	0.475	30	47.0%	49.8%	57.9%
				50	36.4%	47.0%	55.4%
				100	25.7%	44.8%	53.7%
				200	18.2%	43.7%	53.0%
0.9	0.05	0.9	0.135	30	72.1%	51.6%	56.3%
				50	70.1%	48.6%	54.1%
				100	68.0%	46.2%	52.3%
				200	66.5%	44.9%	51.5%
0.9	0.01	0.5	0.059	30	87.3%	72.5%	77.8%
				50	86.3%	70.5%	76.7%
				100	85.4%	69.0%	72.2%
				200	84.7%	68.3%	69.3%
0.9	0	0.05	0.005	30	100.0%	96.6%	81.7%
				50	100.0%	96.4%	74.0%
				100	100.0%	96.2%	67.2%
				200	100.0%	96.2%	63.7%

"Discounts" from sample mean for lower 97.5% confidence intervals for hypothetical data. Of n draws, np give data λ and $n(1-p)$ give data ν. Sample mean is unbiased for the fraction of the upper bound that is attained for the population. Column 7 is average over 100,000 permutations.

Distribution	True Mean	n	Average Discount from True Mean		
				Kaplan	
			MDKW	1-perm	all-perm
$U[0,1]$	0.5	30	42.9%	27.9%	33.1%
		50	34.4%	24.3%	30.4%
		100	25.2%	21.5%	28.4%
		150	20.9%	20.6%	27.7%
		200	18.2%	20.1%	27.4%
		250	16.4%	19.8%	27.2%
$(N(0.5, 0.25) \wedge 1) \vee 0$	0.5	30	39.8%	30.6%	27.0%
		50	32.2%	28.7%	24.3%
		100	23.9%	28.2%	22.3%
		150	20.0%	28.1%	21.8%
		200	17.6%	28.0%	21.5%
		250	16.0%	28.0%	21.4%

Mean percentage discount from the population mean to obtain 97.5% lower confidence bounds n independent draws. In the first population, $\{X_j\}$ are iid $U[0,1]$. In the second population, $\{X_j\}$ are iid truncated normals: $X_j \sim (Z \wedge 1) \vee 0$, where $Z \sim N(0.5, 0.25)$. True population mean is 0.5. The average amount by which the lower confidence bound is below the population mean for each of the three lower confidence bounds is given in columns 4–6. 100,000 replications.

Why not use the Normal Approximation?

# Small Values	Sample Size 30	50	100	150	200	250
1	7.8%	13.0%	23.7%	33.2%	41.8%	49.1%
2	15.1%	23.7%	41.5%	55.7%	66.3%	74.5%
3	21.8%	33.5%	55.8%	70.7%	80.5%	87.1%
4	28.0%	42.0%	66.1%	80.4%	88.8%	93.4%
5	33.7%	49.3%	74.4%	87.0%	93.4%	85.4%
10	56.2%	74.7%	93.7%	91.7%	90.8%	90.8%
15	70.8%	87.5%	91.7%	94.5%	91.1%	94.3%
20	81.0%	93.7%	91.0%	91.4%	91.9%	92.5%
25	87.7%	85.8%	91.1%	94.4%	92.9%	95.5%

Approximate coverage probability of nominal 97.5% lower confidence bounds based on the normal approximation for a population with two distinct values. 370 items in the population. Draws with replacement with equal probability. The population contains two distinct values, v_1 and v_2, with $v_1 < v_2$. Column 1: number in the population equal to v_1, the smaller value. Columns 2–7: empirical percentage of intervals that cover in 100,000 replications.

# Each of Two Smaller Values	Sample Size 30	50	100	150	200	250
1	15.1%	23.8%	41.8%	55.9%	66.0%	74.1%
2	27.9%	41.7%	66.3%	80.4%	76.3%	84.4%
3	39.0%	55.9%	80.4%	80.9%	84.6%	88.0%
4	48.2%	66.6%	76.6%	84.6%	87.6%	90.0%
5	55.7%	74.4%	84.5%	87.9%	89.8%	92.7%
10	81.2%	84.8%	90.3%	93.2%	93.3%	93.6%
15	81.4%	88.6%	93.3%	94.2%	94.4%	94.9%
20	85.8%	90.8%	93.8%	94.5%	94.9%	95.5%
25	92.3%	93.5%	94.3%	95.0%	95.3%	95.7%

Approximate coverage probability of nominal 97.5% lower confidence bounds based on the normal approximation, for a population with three equispaced values. 370 items in the population. Draws with replacement with equal probability. The population contains three distinct equally-spaced values, v_1, v_2 and v_3, where $v_1 < v_2 < v_3$ and $v_2 - v_1 = v_3 - v_2$. Column 1: number in the population equal to each of the two smaller values. v_1 and v_2. Columns 2–7: empirical percentage of intervals that cover in 100,000 replications.

Scale Factor	Sample Size 30	50	100	150	200	250
1	27.8%	35.1%	55.8%	59.4%	68.0%	74.9%
5	62.9%	74.0%	84.8%	88.4%	90.3%	91.4%
10	78.1%	85.2%	90.7%	92.5%	93.3%	94.0%

Approximate coverage probability of nominal 97.5% lower confidence bounds based on the normal approximation, for a population with logistic values. 370 items in the population. Draws are with replacement with equal probability. The values in the population follow a logistic curve: The kth member of the population has value $(1+\exp(-k/s))^{-1}$ for a fixed scale factor s. Column 1: scale factor s. Columns 2–7: empirical percentage of intervals that cover in 100,000 replications of drawing samples of size 30–250.

Novato Sanitary District Director, Contested Election, 2009

Vote for 3 candidates; 30 precincts, 10,945 ballots cast (35.63% turnout, 72.9% VBM) Undervote rate $>$ 13%.

DENNIS J. WELSH	5844	20.47%
MICHAEL DI GIORGIO	4621	16.19%
BILL LONG	4338	15.20%
BILL SCOTT	4323	15.14%
ARTHUR T. KNUTSON	3726	13.05%
DENNIS FISHWICK	3506	12.28%
E.A. SAM RENATI	2161	7.57%
Write-in Votes	29	0.10%
Total Votes	28548	

15-vote margin for Long over Scott.

Election was audited; outcome confirmed to the satisfaction of Scott.

Later discovered that as many as 67 voters didn't get the right ballot, owing to clerical errors in tax offices, etc. (not Registrar of Voters).

Of the 67, only 28–33 voted in election.

Strict reading of California Law Governing Election Contests: Must show convincing evidence that voters tried to vote but were denied the right, and that those denied would have changed the outcome.

Contestants presented no evidence that outcome would be different.

1. Only know of 2–4 who attempted and were denied the opportunity to vote: one voted provisionally, one tried to. Could not change the outcome.

2. The 28–33 voters who didn't get the right ballot all VBM. Margin for Long $\approx 20\times$ for VBM voters than on the whole. (1.16% versus 0.06%).

3. If 33 were a precinct, would require 60% larger margin in favor of Scott (45.45%) than any of the 30 real precincts had (28.18%).

4. If 33 were random sample and outcome really was a tie or a win for Scott, $\mathbb{P}$(15-vote margin hidden among 33) $<$ 0.68%. (values -1, 0, 1. Condition on nonzero; max over conditional probabilities.)

$$
\begin{aligned}
P(X \geq 15) &= P(X \geq 15|\text{no 0s})P(\text{no 0s}) + \\
&\quad +P(X \geq 15|\text{1 0})P(\text{one 0}) + \cdots \\
&\quad +P(X \geq 15|\text{18 0s})P(\text{18 0s}) \\
&\leq \max_{j=0}^{18} P(X \geq 15|j \text{ 0s})P(j \text{ 0s}). \qquad (1)
\end{aligned}
$$

5. Map of wrong ballots shows no pattern. No evidence presented that voters who got got wrong ballot would vote as bloc.

6. Survey: no surprising pattern of votes for Scott versus Long. Results for 6 of the 33 (4 Long, 3 Scott). $\mathbb{P}$(16-vote margin hidden among 27) $< 0.12\%$.

How many would it take to make it plausible that outcome would be different?

Absent evidence that the 33 would have voted differently from the rest, need ≈1000–2300 to have gotten wrong ballots to have moderate chance they would change outcome.

Arguments: sampling calculation, median margin.