Probabilistic Methods in Social Choice

Eric Pacuit
University of Maryland

August 14, 2021

Plan

- ✓ Background on voting theory
- ✓ Generating preference profiles
- ✓ Quantitative analysis of voting methods
- ✓ Probabilistic voting methods
- ⇒ Condorcet jury theorem and related results
- ▶ Aggregating probabilistic judgements

M. Pivato. *Voting rules as statistical estimators*. Social Choice and Welfare, 40:581 - 630, 2013.

$\mathcal{S}$ is a set of states where $s^* \in \mathcal{S}$ is the unknown *true* state.

$\mathcal{I}$ i s a collection of voters

$\mathcal{X}$ is a set of outcomes (alternatives, rankings, judgements, etc.) and $\mathcal{V}$ is a set of signals, or votes.

A profile is an element $\boldsymbol{v} \in \mathcal{V}^{\mathcal{I}}$

A voting method is a function $F : \mathcal{V}^{\mathcal{I}} \rightarrow \wp(\mathcal{X})$ assigning a non-empty subset of $\mathcal{X}$ to each profile $\boldsymbol{v}$

For all $i \in \mathcal{I}$, let $v^i \in \mathcal{V}$ be a message indicating the beliefs of voter i about the true state, i.e., v^i is a 'noisy signal' of s^*.

For $i \in \mathcal{I}$, $s \in \mathcal{S}$ and $v \in \mathcal{V}$, let $\rho^i_s(v)$ be the conditional probability that voter i will send the signal v when the true state is s. An **error model** is a function:

$$\rho : \mathcal{I} \times \mathcal{S} \to \Delta(\mathcal{V})$$

An error model is **anonymous** when $\rho(i, s) = \rho(j, s)$ for all i, j and s. Then we can represent an anonymous error model as $\rho : \mathcal{S} \to \Delta(\mathcal{V})$.

The conditional probability of seeing profile $\boldsymbol{v}$ is $R : \mathcal{S} \to \Delta(\mathcal{V}^{\mathcal{I}})$ defined by

$$R(s; \boldsymbol{v}) = \Pi_{i \in \mathcal{I}} \rho_{\boldsymbol{s}}(i, v^i)$$

Let $\alpha \in \Delta(\mathcal{S})$ be a prior probability density on $\mathcal{S}$. For all $s \in \mathcal{S}$, the posterior probability of s, denoted $\omega_{\boldsymbol{v}}(s)$, is

$$\omega_{\boldsymbol{v}}(s) = \frac{R(s; \boldsymbol{v})\alpha(s)}{\overline{R}(\boldsymbol{v})}, \qquad \text{where } \overline{R}(\boldsymbol{v}) = \int_{\mathcal{S}} R(s; \boldsymbol{v})\alpha(s)\mathrm{d}s$$

The maximum a posteriori (MAP) estimator is the set of all $s \in \mathcal{S}$ which have maximal a posteriori probability:

$$MAP^{\mathcal{S}}_{\alpha,\rho}(\boldsymbol{v}) = \text{argmax}_{s\in\mathcal{S}}\omega_{\boldsymbol{v}}(s) = \text{argmax}_{s\in\mathcal{S}}(R(s;\boldsymbol{v})\alpha(s))$$

If we assume the prior probability α is uniformly distributed over $\mathcal{S}$, thenS, then $MAP^{\mathcal{S}}_{\alpha,\rho}(\boldsymbol{v})$ coincides with the maximum likelihood estimator (MLE), defined as

$$MLE^{\mathcal{S}}_{\rho}(\boldsymbol{v}) = \text{argmax}_{s\in\mathcal{S}}R(s;\boldsymbol{v})$$

The maximum a posteriori (MAP) estimator is the set of all $s \in \mathcal{S}$ which have maximal a posteriori probability:

$$MAP^{\mathcal{S}}_{\alpha,\rho}(\boldsymbol{v}) = \mathrm{argmax}_{s\in\mathcal{S}}\omega_{\boldsymbol{v}}(s) = \mathrm{argmax}_{s\in\mathcal{S}}(R(s;\boldsymbol{v})\alpha(s))$$

If we assume the prior probability α is uniformly distributed over $\mathcal{S}$, thenS, then $MAP^{\mathcal{S}}_{\alpha,\rho}(\boldsymbol{v})$ coincides with the maximum likelihood estimator (MLE), defined as

$$MLE^{\mathcal{S}}_{\rho}(\boldsymbol{v}) = \mathrm{argmax}_{s\in\mathcal{S}}R(s;\boldsymbol{v})$$

The Condorcet Jury Theorem says that the majority voting rule is an MLE when $\mathcal{S} = \mathcal{V} = \{1, -1\}$. Which other voting rules can function as MAP or MLE for other choices of ρ and α?

Assume that $\mathcal{X} = \mathcal{S}$.

A voting rule $F : \mathcal{V}^{\mathcal{I}} \to \mathcal{X}$ is *MAP*-rationalizable if there is a prior probability density $\alpha \in \Delta(\mathcal{X})$ and an error model $\rho : \mathcal{I} \times \mathcal{X} \to \Delta(\mathcal{V})$ such that $F(\boldsymbol{v}) = MAP^{\mathcal{X}}_{\alpha,\rho}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \mathcal{V}^{\mathcal{I}}$.

F is *MLE*-rationalizable if F is *MAP*-rationalizable for the uniform prior probability density function on $\mathcal{X}$.

Theorem (Pivato)

1. F is MAP-rationalizable if and only if F is a scoring rule.
2. F is MLE-rationalizable if and only if F is a *balanced* scoring rule.
3. F is anonymously *MLE*-rationalizable if and only if F is an anonymous, balanced scoring rule.

See, also:

H. P. Young. *Optimal voting rules*. Journal of Economic Perspectives, 9(1):51 - 64, 1995.

V. Conitzer and T. Sandholm. *Common voting rules as maximum likelihood estimators*. In: 21st Annual conference on uncertainty in artificial intelligence (UAI-05), pp 145 - 152, 2005.

L. Xia and V. Conitzer. *A maximum likelihood approach towards aggregating partial orders*. In: 23rd International joint conference on artificial intelligence (IJCAI-11), pp 446 - 451, 2011.

Borda

To *MLE*-rationalize Borda, suppose each voter is most likely to choose a preference order that judges x^* (the "correct outcome") to be best, and less likely to choose a preference order where x^* has lower rank, with probability exponentially decreasing according to the rank of x^*.

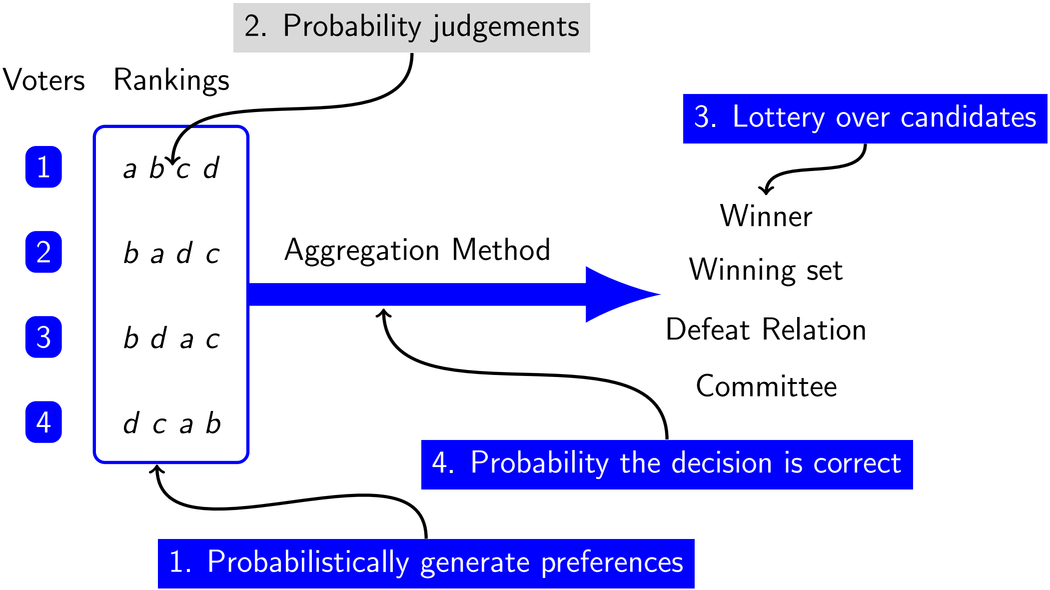
Let $S(v, x)$ be the score of x in the ranking v and let $\epsilon \in (0, 1)$. For all $x \in \mathcal{X}$ and $v \in \mathcal{V}$ (a ranking) suppose that:

$$\rho_x(v) = \frac{\epsilon^{S(v,x)}}{C}, \qquad \text{where} \quad C = (N-1)!\frac{\epsilon^N - 1}{\epsilon - 1}$$

Theorem (Pivato). Let $\mathcal{X}$ be a finite set. Let $\mathcal{V}$ be an arbitrary set. Let $F^* : \mathcal{V}^* \to \mathcal{X}$ be a neutral, anonymous, variable-population voting rule. Then F^* is MLE-rationalizable if and only if F^* satisfies **reinforcement** and **overwhelming majority**.

...the statistical rationalization approach begins with a familiar voting rule, and then contrives some probabilistic scenario to rationalize it as a statistical estimator ex post facto. But this is backwards. One should begin by specifying the scenario which best describes the epistemic problem faced by the voters, and then derive the correct statistical estimator for this scenario.

M. Pivato. *Realizing epistemic democracy*. in The future of economic design (J.-F. Laslier, H. Moulin, R. Sanver, and W. S. Zwicker, eds.), Springer-Verlag 2019, pp. 103 -112.


2. Probability judgements
Voters
Rankings
3. Lottery over candidates
1
a b c d
2
b a d c
3
b d a c
4
d c a b
Aggregation Method
Winner
Winning set
Defeat Relation
Committee
4. Probability the decision is correct
1. Probabilistically generate preferences

Majority Rule

All voting methods reduce to majority rule on 2-candidate profiles.

Proposition

For any VCCR f, the following are equivalent:

1. *f coincides with majority rule on two-candidate profiles;*
2. *f satisfies the following axioms with respect to two-candidate profiles: Anonymity, Neutrality, Monotonicity, Pareto, and Upward Neutral Reversal.*

W. Holliday and EP. *Axioms for Defeat in Democratic Elections*. Forthcoming Journal of Theoretical Politics, `https://arxiv.org/pdf/2008.08451.pdf`, 2021.

Theorem (Smith 1973, Young 1974)

A voting method satisfies Anonymity, Neutrality and **Reinforcement** *if and only if F is a scoring rule.*

Saari's argument, Balinski and Laraki (2010, pg. 77); Zwicker (2016, Proposition 2.5): Multiple districts paradox, f *cancels properly*.

2	2	2	1	2
a	b	c	a	b
b	c	a	b	a
c	a	b	c	c

- no Condorcet winner in the left profile
- b is the Condorcet winner in the right profile
- a is the Condorcet winner in the combined profiles

Condorcet's Other Paradox

# voters	30	1	29	10	10	1
s_2	a	a	b	b	c	c
s_1	b	c	a	c	a	b
s_0	c	b	c	a	b	a

Condorcet's Other Paradox: No *scoring rule* will work...

$Score(a) = s_2 \times 31 + s_1 \times 39 + s_0 \times 11$

$Score(b) = s_2 \times 39 + s_1 \times 31 + s_0 \times 11$

$Score(a) > Score(b) \Rightarrow 31s_2 + 39s_1 > 39s_2 + 31s_1 \Rightarrow s_1 > s_2$

$b >_{BC} a >_{BC} c$ $\qquad$ $a >^M b >^M c$

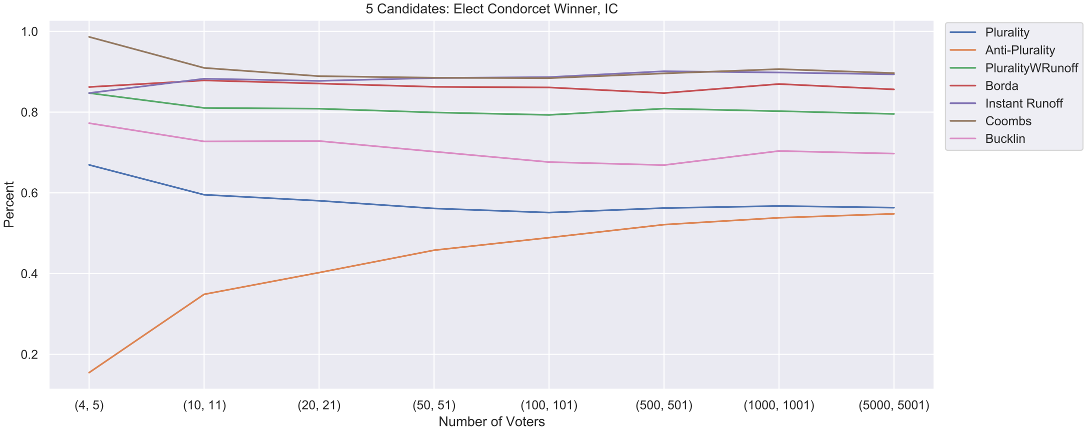
Condorcet's Other Paradox

# voters	30	1	29	10	10	1
s_2	a	a	b	b	c	c
s_1	b	c	a	c	a	b
s_0	c	b	c	a	b	a

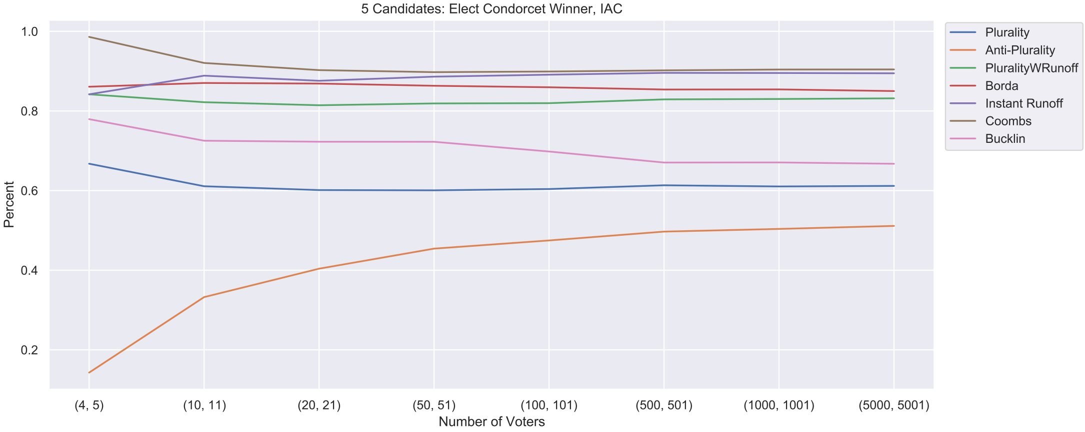
Theorem (Fishburn 1974). For all $m \geq 3$, there is some voting situation with a Condorcet winner such that every scoring rule will have at least $m - 2$ candidates with a greater score than the Condorcet winner.

P. Fishburn. *Paradoxes of Voting*. The American Political Science Review, 68:2, pgs. 537 - 546, 1974.

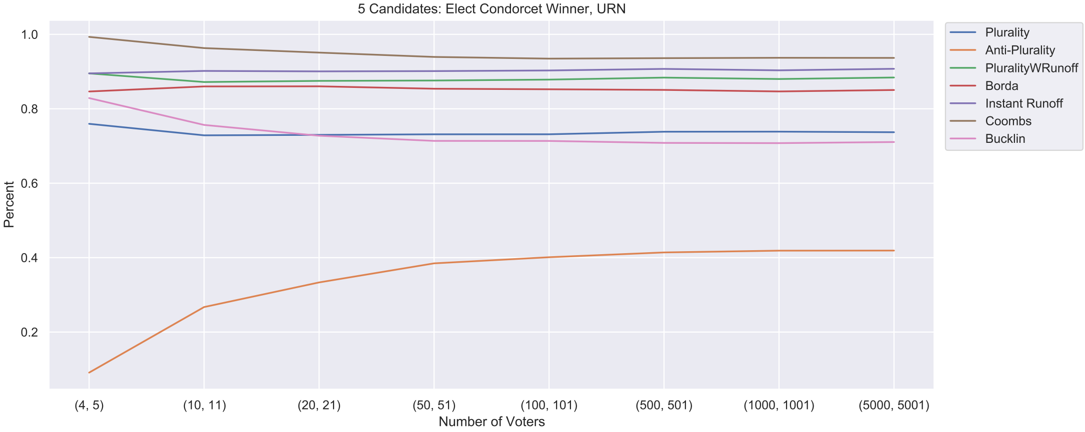
Condorcet Efficiency



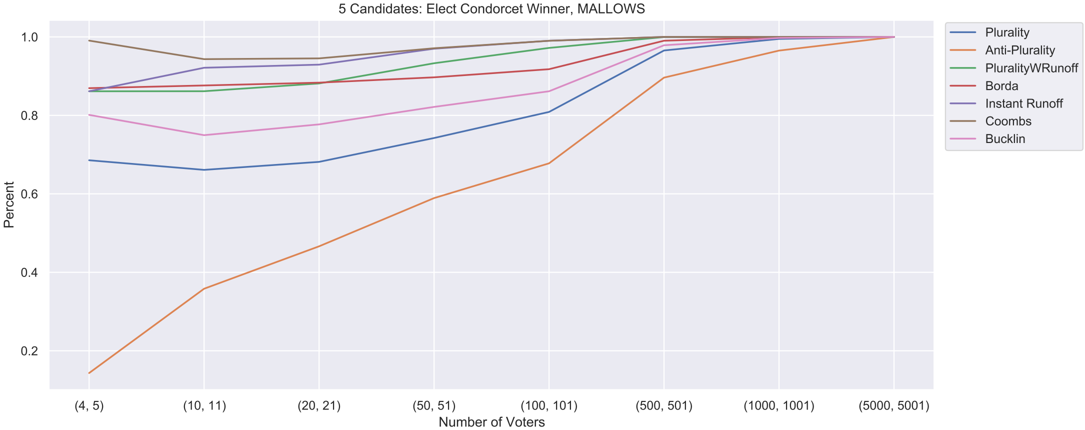
Condorcet Efficiency



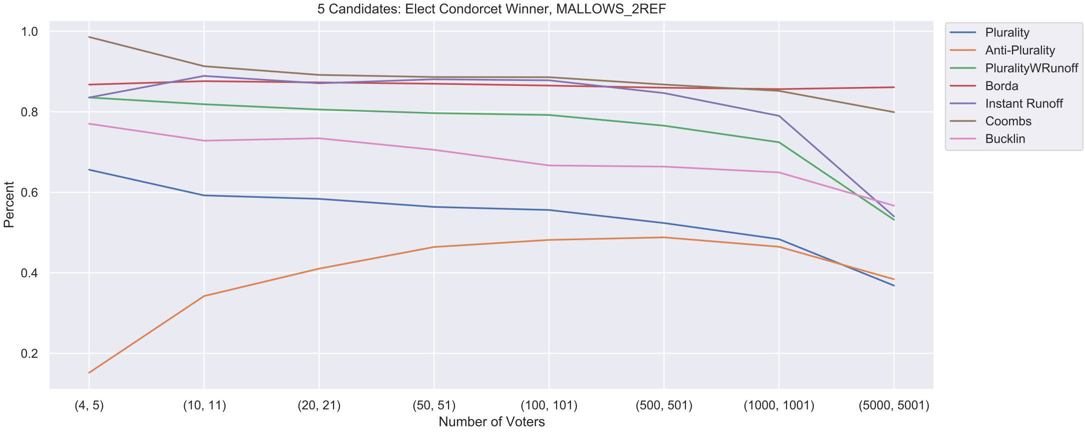
Condorcet Efficiency



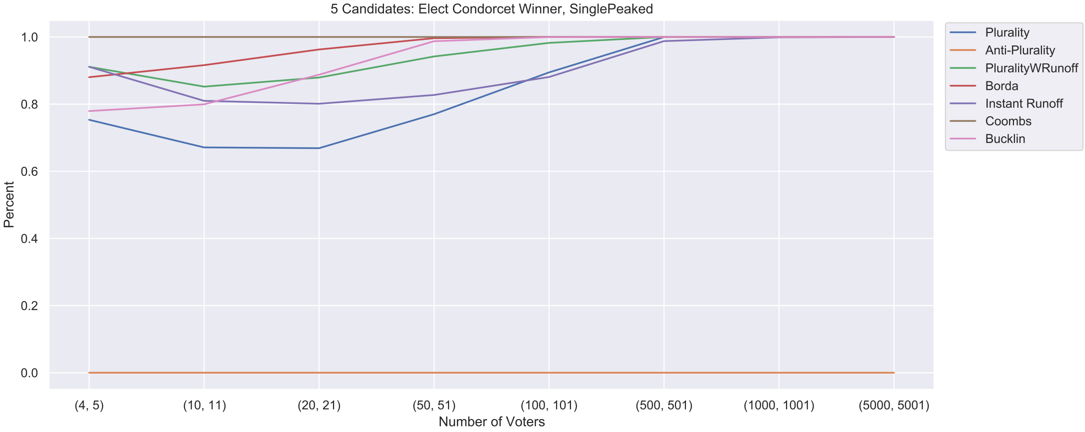
Condorcet Efficiency



Condorcet Efficiency



Condorcet Efficiency



Theorem (Brandl et al., 2016) *ML* is the only anonymous PSCF satisfying population-consistency, cloning-consistency, and Condorcet-consistency when preferences are strict.

Population-consistency: whenever two disjoint electorates agree on a lottery, this lottery should also be chosen by the union of both electorates (aka. reinforcement).

Cloning-consistency: the probability that an alternative receives is unaffected by introducing new variants of another alternative. Alternatives are variants of each other if they bear the same relationship to all other alternatives and therefore constitute a contiguous interval in each voter's preference ranking.

F. Brandl, F. Brandt, and H. G. Seedig. *Consistent probabilistic social choice*. Econometrica, 84(5), pgs. 1839 - 1880, 2016.

We are interested in voting methods that:

- ✓ respond in a reasonable way to **new candidates** joining the election (Stability for Winners, Immunity of Spoilers);
- ✓ respond in a reasonable way to **new voters** joining the election (Positive Involvement, Negative Involvement).

	Split Cycle	Ranked Pairs	Beat Path	Mini-max	Copeland	Borda	Coombs	Instant Runoff	Plurality
Condorcet Winner	✓	✓	✓	✓	✓	–	–	–	–
Condorcet Loser	✓	✓	✓	–	✓	✓	✓	✓	–
Monotonicity	✓	✓	✓	✓	✓	✓	–	–	✓
Immunity to Spoilers	✓	–	–	✓	✓	–	–	–	–
Stability for Winners	✓	–	–	–	–	–	–	–	–
Positive Involvement	✓	–	–	✓	–	✓	–	✓	✓
Negative Involvement	✓	–	–	✓	–	✓	✓	–	✓

Thank you!