A statistical approach to epistemic democracy

Marcus Pivato
Trent University
June 28, 2011

Abstract. We briefly review Condorcet and Young's epistemic interpretations of preference aggregation rules as maximum likelihood estimators. We then develop a general framework for interpreting epistemic social choice rules as maximum likelihood estimators, maximum *a posteriori* estimators, or expected utility maximizers. We illustrate this framework with several examples. Finally, we critique this program.

Consider a group of voters trying to collectively determine the correct answer to some factual question, which has an objectively correct (but unknown) answer. Suppose all the voters have the same values or preferences; the only conflict is over their beliefs about objective facts.[1] What is the best way for the voters to reconcile their contradictory beliefs and arrive at a collective decision? We can distinguish between three sorts of questions:

1. A question about the best policy or action to take. (e.g: “What action (if any) should society take to counter anthropogenic climate change?”)
2. A question about a matter of fact, where there is some obvious background probability distribution over the possible answers (e.g. “What will the weather be in Pittsburgh on June 28?” Here, the background distribution could be based on historical data about the weather on June 28 in previous years.)
3. A question about a matter of fact, where there is no obvious background probability distribution. (e.g. “What caused the Permian-Triassic extinction event?”)

In Type 1 questions, the word “best” suggests that the group wishes to maximize the value of some (universally agreed upon) utility function. When confronting uncertainty (where each action yields a probability distribution over the set of outcomes), the standard approach[2] is to choose the action(s) with the largest *expected* utility. In other words, the social decision should be an *expected utility maximizer* (EUM).

In questions of Type 2 or 3, there is no utility function we wish to maximize ---either because the question is of purely “academic” interest (e.g. paleontology), or because different people may have different utility functions (e.g. weather), and the group as a whole does not wish to fixate on any

1 Cohen (1986) and Estlund (1997) describe this kind of `purely epistemic' social decision problem as `epistemic democracy'. This model is in direct contrast to much of social choice theory, which assumes that voters agree about the objective facts, but have irreconcilable differences in their preferences or values.

2 By “standard”, I mean that this approach is almost universally accepted in economics, finance, and management science. Furthermore, it is the unique approach which satisfies certain axioms of “rationality” or “consistency”, as demonstrated by the well-known theorems of von Neumann and Morgenstern, Savage, etc. Of course, there are other (“non-expected utility”) approaches to risky decision-making; these may be better descriptions of actual human behaviour, but they are hard to defend on the grounds of “rationality”. Also, of course expected-utility maximization requires the specification of a probability distributions associated with each action; in some situations, these may be unknown or unknowable.

particular utility function. In a Type 2 question, our background knowledge can be represented by a *prior probability distribution* over the possible answers; society's goal is to combine this prior distribution with the information provided by the voters to identify the answer which is most likely to be correct, all things considered. If there is a well-defined probabilistic relationship between the possible answers and the possible signals sent by voters, then the correct way to do this is to use Bayes rule to compute a *posterior probability distribution,* and then choose the answer(s) with the largest posterior probability. In the jargon of statistics, the social decision should be a *maximum a posteriori* (MAP) estimator. In a Type 3 question, there is no natural prior distribution, and all the possible answers are assumed to be equally likely, *a priori*.[3] In this situation, we begin with the uniformly distributed prior distribution, and again apply Bayes rule to identify the most likely outcome; thus, the social decision will be a *maximum likelihood estimator* (MLE).

Condorcet (1785) was the first to suggest that a social decision could be an MLE; the celebrated *Condorcet Jury Theorem* (CJT) says that, under certain hypotheses, simple majority vote is an MLE when society must answer a yes/no question. Condorcet's hypotheses (that the voters are independent and identically distributed random variables) were highly unrealistic; there are now many variations of the CJT with more realistic hypotheses.[4] Meanwhile Young (1986,1988,1995,1997) has shown that the Kemeny rule and the Borda rule can both be interpreted as MLEs, when society faces a preference-aggregation problem involving more than two alternatives. Conitzer, Xia, and their collaborators have extended the Condorcet-Young approach, and asked what other preference aggregation rules can be interpreted as MLEs.[5]

However, preference aggregation is only one kind of social choice problem, and the MLE is only one of the three types of epistemic rules described above. What other voting rules can be interpreted as MLE, MAP, or EUM procedures, given suitable assumptions about the nature of the decision problem and the background knowledge of the voters? This paper will first describe the main results of Pivato (2011), which offer a partial answer to this question. Then we will critique this program.

Let $I=\{1,2,3,...,N\}$ represent a set of voters, and let V be the set of signals (i.e. "votes") which could be sent by each voter. A *profile* is a list $\mathbf{v}$ which assigns a signal $v_i \in V$ to each voter $i \in I$. Let V^I denote the set of all profiles, and let X be a set of alternatives available to society (e.g. possible actions, possible answers to some question). A *voting rule* is a correspondence F from V^I to X ---that is, for any profile $\mathbf{v} \in V^I$, we obtain a nonempty subset $F(\mathbf{v}) \subseteq X$. Typically (but not always) $F(\mathbf{v})$ will be a singleton.

For example, fix a function S from $I \times V \times X$ into $\mathbb{R}$.[6] For any $\mathbf{v} \in V^I$, let $F_S(\mathbf{v})$ be the set of all element(s) $x \in X$ which maximize the sum $S(1,v_1,x)+S(2,v_2,x)+....+S(N,v_N,x)$. The resulting correspondence F_S from V^I into X is the *scoring rule* defined by S. Many common voting rules can be

3 This so-called *Principle of insufficient reason* (PIR) is only really defensible when no obvious asymmetry exists between different possible answers, so they are effectively interchangeable. Furthermore, if the set of answers is infinite, then it is impossible for them to all be "equally probable". In this case, the PIR is only meaningful when there is some underlying "canonical" probability measure, such as the Lebesgue measure. Even here, if the space of answers is unbounded, the Lebesgue measure cannot generally be "normalized" to a probability measure. This is why I use the vague phrase "equally likely", rather than "equally probable". The fact remains that the PIR ---and, thus MLE ---lacks a satisfactory conceptual foundation. The main argument in its favour is the absence of any obviously better alternative.

4 See, for example, List and Goodin (2001), Nitzan (2010, Part III), Hummel (2010), or Dietrich and Spiekerman (2011).

5 See Conitzer and Sandholm (2005), Conitzer et al. (2009), Xia et al (2010), and Conitzer and Xia (2011).

6 Here, $I \times V \times X$ is the set of all ordered triples (i,v,x), where $i \in I$, $v \in V$, and $x \in X$. Meanwhile, $\mathbb{R}$ is the set of real numbers.

represented as scoring rules. For example:

- In the *plurality rule*, $V=X$, and $S(i,v,x)=1$ if $v=x$, whereas $S(i,v,x)=0$ if $v\neq x$.
- In the *weighted plurality rule*, $V=X$, and $S(i,v,x)=w_i$ if $v=x$, whereas $S(i,v,x)=0$ if $v\neq x$. Here, w_i is the "weight" of voter *i,* which perhaps is a measure of her estimated "competency".
- In the *antiplurality rule*, $V=X$, and $S(i,v,x)=-1$ if $v=x$, whereas $S(i,v,x)=0$ if $v\neq x$ (i.e. each voter votes *against* her least-prefered alternative, implicitly endorsing all the other alternatives).
- In *approval voting,* V is the set of all subsets of X, and $S(i,v,x)=1$ if $x\in v$, whereas $S(i,v,x)=0$ if $x\notin v$ (Brams and Fishburn, 1983).
- In *range voting*, V is the set of all functions mapping each element of X into a real number between 0 and 1, and $S(i,v,x)=v(x)$ (Smith, 2001).
- In the *Borda rule,* V is the set of all strict rankings of $X,$ and $S(i,v,x)=r$ if x is ranked rth from the bottom according to the ranking v.
- In the *Kemeny* (1959) *rule*, both V and X are the set of all strict rankings over some set A of alternatives, and $S(i,v,x)$ is the number of pairwise orderings where v and x agree.

Note that, aside from the weighted plurality rule, all of these scoring rules are *anonymous*, in the sense that $S(i,v,x)=S(j,v,x)$ for all $i,j\in I$, $v\in V$, and $x\in X$. In other words, all voters have exactly the same "weight". If $S(i,v,x)=-\infty$ for some $i\in I$, $v\in V$, and $x\in X$, then voter i can effectively "veto" the choice x by sending the signal v. A rule has *no vetos* if this is never the case. All the above rules have no vetos.

Now suppose that X represents a set of possible "states of nature"; the true state is unknown. We suppose that each voter receives some partial information about the true state, which determines the way she votes. We can mathematically model these assumptions by specifying a prior probability distribution over X, and, for each $i\in I$ and $x\in X$, a probability distribution over V (called the *error model*), which describes the sort of signal which voter i is likely to send if the true state of nature is x. (We assume that the signals of different voters are conditionally independent random variables, for any state of nature). We call a combination of a prior and an error model a *scenario*. We say this scenario is *anonymous* if all voters are equally competent, and receive the same quantity and quality of information (i.e. for any $x\in X$, we have the same probability distribution on V for every $i\in I$). If the conditional probability of voter i sending signal v is *zero*, given state of nature x, then the scenario says certain events are "impossible". If this never occurs, for any $i\in I$, $v\in V$, and $x\in X$, then we say the scenario has *no impossibilities*.

Given any scenario C, and any profile $\mathbf{v}\in V^I$, we can use Bayes rule to compute the posterior distribution over X, conditional on $\mathbf{v}$. Let MAP(C,$\mathbf{v}$) denote the element(s) of X which obtain maximal probability in this posterior distribution (the *maximum a posteriori estimator* determined by C and $\mathbf{v}$). A voting rule F is *MAP-rationalizable* if there exists some scenario C such that $F(\mathbf{v}) = \text{MAP}(C,\mathbf{v})$ for all $\mathbf{v}\in V^I$. The rule F is *anonymously MAP-rationalizable* if the scenario C is anonymous. We now come to one of our main results.

Theorem 1: *A voting rule is MAP-rationalizable if and only if it is a scoring rule. Furthermore, it is anonymously MAP-rationalizable if and only if it is an anonymous scoring rule. Finally, it has no vetos if and only if the corresponding scenario has no impossibilities*[7].

7 Pivato (2011, Theorem 1.1).

In particular, the (weighted) (anti)plurality, Borda, Kemeny, approval, and range-voting rules are all MAP-rationalizable. A MAP-rationalizable voting rule F is *MLE-rationalizable* if the rationalizing scenario C has a uniformly distributed prior probability distribution. This is the case if and only if the scoring function S satisfies a technical condition called "balance", which means that all alternatives in X are treated roughly the same.

Theorem 1 tells us that many voting rules (i.e. non-scoring rules) can have *no* epistemic rationalization. However, the fact that a given voting rule *can* be rationalized by some scenario does not imply that this scenario is a particularly realistic description of the epistemic problem facing society. What would a realistic description look like? Typically, the space X of alternatives has a sort of "geometry"; there is a sense in which some elements of X are "close together" (i.e. very similar), whereas other elements are "far apart" (very dissimilar).[8] Suppose each voter's signal is a guess about the correct element in X (so in this model, $V=X$). If the voters are not completely unreliable, then they are more likely to guess an answer which is close to the right answer than one which is far away. Thus, if the true state of nature is x, then we expect there to be some (decreasing) real-valued function E_i such that, for any $v \in X$, the probability that voter i guesses v is given by $E_i[d(x,v)]$ (where $d(x,v)$ is the "distance" between x and v in X). This is called a *metric* error model. In this case, the scoring rule described in Theorem 1 is a *metric* voting rule: for each $i \in I$ and $x,v \in X$, the value of $S(i,v,x)$ is a decreasing function of the distance between v and x. In this setup, the prior probability distribution on X translates into a "bias function" β,[9] and the rule chooses the x in X which maximize the sum

$$\beta(x)+S(1,v_1,x)+S(2,v_2,x)+....+S(N,v_N,x).$$

For example, suppose that the function E_i is *exponentially decaying* (i.e. $E_i(r) = a_i/b_i^r$, for some constants $a_i,b_i>0$). Then it can be shown that (for a suitable prior probability distribution) the MAP estimator is the *weighted median rule.*[10] For any profile **v**, the weighted median rule picks the element(s) of X which *minimize* the sum $w_1\, d(x,v_1) + w_2\, d(x,v_2) + ... + w_N\, d(x,v_N)$, where $w_1,w_2,....,w_N$ are nonnegative "weights" assigned to the voters. The rule is anonymous if and only if $a_1=a_2=....=a_N$ and $b_1=b_2=....=b_N$, or equivalently, $w_1=w_2=....=w_N=1$; this yields the (unweighted) *median rule* (which minimizes $d(x,v_1) + ... + d(x,v_N)$). The space X is called *homogeneous* if the geometry of X "looks the same" around each x in X.[11] In this case, the median rule is the MLE for any anonymous exponential error model (where $E_i(r) = a/b^r$, for all i in I, for some constants $a,b>0$ which are independent of i).

For example, let X be the space of all strict preference relations on some set A of alternatives. For any x and y in X, let $d(x,y)$ be the number of pairs in A on which the orders of x and y disagree. Then (X,d) is a homogeneous space. The median rule on X is the *Kemeny rule*. The argument sketched above shows that the Kemeny rule is the MLE for any anonymous exponential error model on X; this was first noted by Young (1986,1988,1995,1997).

---

8 Formally, this means X is a *metric space*.

9 Note that a uniform prior does not necessarily translate into a constant bias function. The bias is generated by an interaction between the prior probability distribution, the error functions E_i, and the geometry of the metric on X. See Pivato (2011, Theorem 2.1).

10 See Pivato (2011, Section 3).

11 Formally: for any x and y in X, there is an isometry of X which maps x into y. (An *isometry* is a bijection from X to itself which preserves distances.) See Pivato (2011, Corollary 2.2).

Here is another example. Let A be a set of possible candidates for some committee. Let X be the set of all committees comprised of exactly n of these candidates (for some fixed $n \leq |A|$). For any x and y in X, let $d(x,y)$ be the number of candidates on which x and y disagree. Then (X,d) is a homogeneous space; thus, again, the median rule is the MLE for any anonymous exponential error model on X. Likewise, if X is the set of all committees with an *odd* number of candidates, then it is homogeneous, so the median is the MLE.[12]

The median rule is not the only voting rule with interesting statistical rationalizations. Let X be a homogeneous subset of a Euclidean space (i.e. $\mathbb{R}^M$, for some $M>0$), which we endow with the standard Euclidean notion of distance, and consider an anonymous *Gaussian* error model on X (that is: $E_i(r) = \exp(-r^2/2\sigma^2)/C$ for all i in I, for some constants σ, $C > 0$ which are independent of i). The MLE for this error model is the metric voting rule where $S(i,v,x) = -d(x,v)^2$ for all $i \in I$, $v \in V$, and $x \in X$. It is easy to check that this rule simply chooses the element(s) of X which are closest to the *average* of the votes $v_1, v_2, \ldots, v_N$, when treated as M-dimensional vectors in $\mathbb{R}^M$.[13]

For example, let A be a set of N social alternatives. A *ranking* of A is a bijection from A into the set $\{1,2,\ldots,N\}$; we can regard such a ranking as a vector in the Euclidean space $\mathbb{R}^A$. Let X be the set of all such rankings, regarded as a subset of $\mathbb{R}^A$. Then X is homogeneous, and the averaging rule (i.e. the MLE for any Gaussian error model) is the *Borda rule.* This provides another MLE-rationalization for the Borda rule, which is quite different from the rationalization given by Young (1986). Variations of this setup yield MLE-rationalizations of other "ranking" rules, including the plurality rule and anti-plurality rule.

Now let us turn to the problem of expected utility maximization. Many voting rules invite each voter to assign a numerical "score" to each social alternative, and then choose the alternative with the highest average score. Rules differ on what range of scores are admissible. For example, *classical utilitarianism* allows the score of an alternative to be any real number, whereas *range voting* allows any score in the interval [0,1], and *approval voting* requires the score of each alternative to be either 0 or 1. In all these rules, a voter can give the same score to two or more alternatives (unlike the Borda rule, (anti)plurality rule, and other "ranking" rules). We will refer to rules with this property as *quasiutilitarian*.

Any quasiutilitarian voting rule can be rationalized as an expected utility maximizer. Suppose that the score a voter assigns to an alternative reflects her (possibly incorrect) estimate of the "social utility" of that alternative. Although the voters' estimates may be incorrect, we assume that they are not totally perverse: under reasonable assumptions about the voters' (mis)perceptions, one can show that, if alternative a receives an average score of r, and alternative b receives an average score of s, and $r > s$, then the *conditionally expected utility* of a, given this information, is higher than the conditionally expected utility of b. (This does *not* mean that r itself is a good estimate of the expected utility of a ---but this isn't necessary). Under these conditions, it is easy to see that the quasiutilitarian rule in question is an expected utility maximizer. In particular, this yields EUM-rationalizations of classical utilitarianism, range voting, approval voting, and a "relaxed" version of the Borda rule (where a voter can given the same rank to two or more alternatives).[14]

12 Pivato (2011, Example 3.5)
13 Pivato (2011, Proposition 2.5)
14 Pivato (2011, Section 4).

* * *

As these results illustrate, a large variety of voting rules can be "rationalized" as MAP, MLE, or EUM rules for some prior distribution and some more-or-less plausible model of the voters' error patterns. However, there are some foundational problems with this entire program. First of all, this program begins with a familiar voting rule, and then "rationalizes" it with some probabilistic scenario, after the fact. But this is the *reverse* of the correct procedure. One should begin by specifying a prior probability distribution and an error model for the voters which captures the underlying epistemic problem as realistically as possible, and then compute the MLE/MAP/EUM for this model; this may or may not end up being a familiar voting rule.

Second of all, these results (as well as the aforementioned work by Young, Conitzer, Xia, and others) assume that the mistakes made by different voters are *independent* random variables. But this is totally unrealistic, because voters presumably come from similar cultural and educational backgrounds, draw upon the same body of public (mis)information, and are in constant communication with one another. Furthermore, a psychological desire to conform or avoid conflict may lead to "herding" or "groupthink". Lorenz *et al.* (2011) have shown empirically that such "social influences" can undermine the reliability of epistemic social choice mechanisms. Thus, any plausible error model must involve correlated errors. But it is not clear how to realistically model such correlations, and a realistic model might be mathematically intractable. Dietrich and Spiekerman (2011) have extended the CJT to a model where voters draw upon common information sources. It would be fruitful to extend their approach to other epistemic social choice problems.

A particularly intractable form of "error correlation" can arise from *strategic dishonesty* on the part of the voters. Even if all the voters have the same objectives, they may have incentives to exaggerate their views or suppress countervailing evidence to counteract what they believe to be the misperceptions of the other ("sincere but misguided") voters. This was demonstrated mathematically by Austen-Smith and Banks (1996). For an empirical example, one has only to look at the recent debates in climate science or macroeconomics, where apparently sincere scientists (who claim to share similar concerns over the long-term welfare of humanity) sometimes accuse one another of overstating or understating their conclusions.[15]

Finally, the "statistical" approach to epistemic democracy assumes that it is possible to specify, with reasonable accuracy, the prior probability distribution and the error model of the voters. But this is not realistic. Consider the aforementioned debate over climate change. Supposing we wanted to obtain a consensus using the statistical techniques described in this paper. What is a reasonable prior? And how could we even begin to specify an error model for the opinions of the climate scientists?[16]

For some problems, a precise specification may not be necessary. For example, the CJT and the aforementioned EUM-rationalizations of quasiutilitarian rules are each valid for a large class of scenarios. But the MAP-rationalization described by Theorem 1 can often be quite sensitive to the

15 There are, of course, also many participants in these debates who are *not* sincere, or not even real scientists, and who may reasonably be accused of short-sightedness, irrationality and/or outright selfishness. However, considerable disagreement remains even after the obvious propagandists, sycophants, ideologues, and crackpots have been eliminated from the dataset.

16 An attempt to do so would arguably be even more vulnerable to error or bias than the original climate science.

underlying scenario. Perhaps, then we should not seek a voting rule which is the "optimal" statistical estimator for one particular scenario, but rather, seek a rule which is "reasonably reliable" for a broad spectrum of scenarios.

However, to consider even a "spectrum" of scenarios, we must first specify the underlying space of possible answers, and in some cases, it is no obviously correct way to do this. For example, consider a debate amongst scientists over the correct model of some phenomenon (e.g. climate). It is far from clear how to specify the "space of possible scientific models". Which structural parameters should be held fixed, and which should be allowed to vary? And how should these variables be coordinatized? There is no obvious answer to these questions, and the outcome of an MLE or MAP rule will be highly sensitive to how they are answered. In effect, the "social consensus" would be an artifact of how we framed the debate.

Fortunately, the scientific community does not generally resolve dissensus by "voting". Instead, scientists deliberate, scrutinize each other's theories, and zero in on those lacunae in the empirical data which allow the dissensus to even exist. They then seek to fill these lacunae as efficiently as possible.[17] This deliberation and empirical exploration continues until little or no dissensus remains, at which point the use of an epistemic social choice mechanism is no longer necessary.

References

Austen-Smith, D., Banks, J., 1996. Information aggregation, rationality, and the Condorcet jury theorem. American Political Science Review, 34-45.

Bohman, J., Rehg, W. (Eds.), 1997. Deliberative Democracy : Essays on Reason and Politics. MIT Press.

Brams, S. J., Fishburn, P. C., 1983. Approval voting. Birkhauser Boston, Mass.

Cohen, J., 1986. An epistemic conception of democracy. Ethics 97 (1), 26-38.

Condorcet, M. d., 1785. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris.

Conitzer, V., Rognlie, M., Xia, L., 2009. Preference functions that score rankings and maximum likelihood estimation. In: 21st International Joint Conference on Artificial Intelligence (IJCAI-09). Pasadena, CA, pp. 109-115.

Conitzer, V., Sandholm, T., 2005. Common voting rules as maximum likelihood estimators. In: 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI-05). Edinburgh, pp. 145-152.

Dietrich, F., Spiekerman, K., 2011. Epistemic democracy with defensible premises. (preprint).

Estlund, D., 1997. Beyond fairness and deliberation: The epistemic dimension of democratic authority. In: Bohman and Rehg (1997), pp. 173-204.

Hummel, P., 2010. Jury theorems with multiple alternatives. Soc. Choice Welf. 34 (1), 65-103.

Kaniovski, S., 2010. Aggregation of correlated votes and Condorcet's jury theorem. Theory and Decision 69 (3), 453-468.

17 For example, through "crucial experiments", which are designed to unambiguously decide between two competing theories.

Kemeny, J. G., Fall 1959. Math without numbers. Daedalus 88, 571-591.

List, C., Goodin, R. E., 2001. Epistemic democracy: Generalizing the Condorcet Jury Theorem. Journal of Political Philosophy 9 (3), 277-306.

Lorenz, Jan, Rauhut, Heiko, Schweitzer, Frank and Helbing, Dirk. How social influence can undermine the wisdom of crowd effect . Proceedings of the National Academy of Science. May 16, 2011, doi: 10.1073/pnas.1008636108.

Pivato, M., 2011. Voting rules as statistical estimators (preprint).

Smith, Warren D., 2001. Range Voting (unpublished). https://math.cst.temple.edu/~wds/homepage/rangevote.pdf

Xia, L., Conitzer, V., 2011. A maximum likelihood approach towards aggregating partial orders. In: 23rd International Joint Conference on Artificial Intelligence (IJCAI-11). p. (to appear).

Xia, L., Conitzer, V., Lang, J., 2010. Aggregating preferences in multi-issue domains by using maximum likelihood estimators. In: 9th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS-10). Toronto, pp. 399-406.

Young, H. P., 1986. Optimal ranking and choice from pairwise decisions. In: Grofman, B., Owen, G. (Eds.), Information pooling and group decision making. JAI press, Greenwich, CT, pp. 113-122.

Young, H. P., 1988. Condorcet's theory of voting. American Political Science Review 82 (4), 1231-1244.

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

Young, H. P., 1997. Group choice and individual judgments. In: Mueller, D. C. (Ed.), Perspectives on public choice: A handbook. Cambridge UP, Cambridge, UK, Ch. 9, pp. 181-200.