A Probabilistic Model of Social Choice

Michael D. Intriligator [2]

Handout* for Seminar Economics and Computation

Weile Weng
Department of Mathematics
Technical University of Munich

June 2018

"... (economics is) an attempt to discover uniformities in a certain part of reality and not the drawing of logical consequences from a certain set of assumptions regardless of their relevance to actuality. Simplified theory-building is an absolute necessity for empirical analysis; but it is a means, not an end."

— Kenneth J. Arrow, *Social Choice and Individual Values*

*This handout is written heavily based on the Intriligator's paper (see [2]), with comments from other literature and some personal remarks in footnotes. In addition, I expanded the mathematical formulation for axioms, rules and conditions in the main text, with the intention of helping interested reader to understand and/or recall the related property precisely. Please let me know if there shall be any improvement.

Abstract

A probabilistic approach to the problem of social choice is proposed, where the actual choice is made by a random device using social probabilities derived from individual probabilities. Under certain postulated axioms the social probabilities are arithmetic average of the individual probabilities, a result equivalent to assigning equal probabilities to selecting one individual to act as dictator. In contrast to rival rules, such as majority rule, the Borda rule, the Pareto rule, and the intersection rule, the average rule satisfies many desirable conditions and is the only rule that takes consideration of the strength of individual preference.

1 Model

Consider a society consists of m **individuals**(indexed $i = 1, 2, .., m; m \geq 2$), which must choose one of n **alternatives** (indexed $j = 1, 2, ..., n; n \geq 2$), labelled as $A_1, A_2, .., A_n$. Each individual has certain preferences among the alternatives, which are summarized by an individual probability vector. An alternative is generated randomly as the choice of society based on the social preference, summarized as social probability vector. It is generally assumed that individuals have independent preferences, i.e. the individual probability vectors are independent from each other. Formally, one can define the related variables in following way:

- **probability vector space**: $P = \{(p_1, .., p_n) \in [0, 1]^n : \sum_{j \in [n]} p_j = 1\}$
- $\forall i \in [m]$, the i-th **individual probability vector**: $q_i = (q_{i1}, ..., q_{in}) \in P$
- **probability matrix**: $Q = (q_{ij})_{i \in [m], j \in [n]} \in P^m$
- **social probability vector**: $p = (p_1, ..., p_n) \in P$
- **social probability function**: $f : P^m \to P, Q \mapsto f(Q)$

Interpretation:

- *from probability vector to preference*: the probability vector of a subject (individual or society) can be considered the relative frequencies that each of the alternatives would be chosen[1]. Based on the probabilities, one can define a preference relation of alternatives for this subject: for individual i , $A_j \succ_i A_k$ if and only if $q_{ij} > q_{ik}$; $A_j \sim_i A_k$ if and only if $q_{ij} = q_{ik}$; for the society, $A_j \succ A_k$, if and only if $p_j > p_k$, $A_j \sim A_k$ if and only if $p_j = p_k$.
- *from social probability to social choice*: the social choice is a random output from the distribution given by the social probability vector.

[1]Fishburn commented in [1] that the use of choice-probability vector as input data is refreshing, since an individual "may be uncertain as to what he would choose if acting alone, or that his choice depends on stochastic aspects of his environment such as might underlie a random utility or a random preference model".

Interest: find social probability function f that satisfies certain desired conditions.

2 The average rule

There are lots of axioms one can use to derive a social probability function f, aka *rule*. The author chooses following axioms:

(i) *Existence of social probabilities.* For a rule f, given any $Q \in P^m$, there exists unique $p \in P$ such that $f(Q) = p$. In other word, f is well-defined on P^m.

(ii) *Unanimity preserving for a loser.* If all individuals reject an alternative with certainty so does society, i.e. if $q_{ij_0} = 0$ for all i then $p_{j_0} = 0$. It implies that if all individuals choose an alternative with certainty then society will do likewise, i.e. if $q_{ij_1} = 1$ for all i then $p_{j_1} = 1$.

(iii) *Independence.* The social probability of choosing one alternative depends only on the individual probabilities of choosing that alternative, which means for $\forall j \in [n]$, p_j is a function of $(q_{1j}, ..., q_{mj})$.[2]

(iv) *Strict and equal sensitivity to individual probabilities.* Social probabilities are strictly sensitive to individual probabilities in that an increase/decrease in the probability that any one individual will choose a particular alternative[3] always increases/decreases the probability that society will choose this alternative. Furthermore, the social probabilities are equally sensitive to individual probabilities in that equal changes in probabilities by different individuals lead to equal changes in social probabilities.[4]

Thus, axioms (iii) and (iv) would give us following equations which hold for all i, j:

$$\begin{aligned} \frac{\partial p_j}{\partial q_{is}} &= \mu_j > 0 \quad \text{for } s = j \\ &= 0 \qquad\quad\, \text{for } s \neq j \end{aligned} \tag{1}$$

[2] In the original paper, this axiom is not used as sufficient condition for average rule, rather as necessary condition in the later part of the paper. However, without this condition, the other three axioms alone are not enough to derive the average rule, as suggested in [3]. One way to amend this is to add the *independence* condition. Rice suggested an alternative approach in [3], by introducing the notion of *(relative) power of individual to the society*. Interested reader may refer to [3] for more detail.

[3] The reader should notice that given any individual vector one can not change one single probability while the others remain the same, since the sum has to be constant 1.

[4] This implies that the effect of change of some individual probability vector only depends on the difference vector. The effect of change does not depend on the individual probability vector prior to change. However, this axiom does not imply that equal changes in probabilities in different alternatives (column vectors in Q) lead to the same change in their social probabilities. Equal sensibility to alternatives implies permutation invariance of alternatives.

Theorem. *(i) - (iv) determine an unique rule, referred to as the* ***average rule****:*

$$p_j = \frac{1}{m} \sum_{i \in [m]} q_{ij}, \text{ for all } j, \tag{2}$$

i.e. $f(Q) = \mathbf{1}/m \cdot Q$ *, with* $\mathbf{1}$ *the* 1*'s (row) vector of size* m *.*

Proof. From equation (1) one has

$$p_j = C_j + \mu_j \sum_{i \in [m]} q_{ij}$$

By axiom (ii) it follows that $C_j = 0$ and $\mu_j = 1/m$, since $0 = C_j + 0$, and $1 = \mu_j \sum_{i=1}^{m} 1$. Thus, we have the rule in form of equation (2). $\square$

Remark. The average rule is equivalent to selecting one of the individuals and generating random social choice based on her probability vector.

Quiz: Can you think of an example that is not average rule but satisfies all axioms except *independence*?

3 Some rival rules and their modifications

Besides the *average rule*, there are other rules that help to determine social probabilities from individual probabilities:

A. **Majority rule**: an alternative is chosen with certainty if a majority prefers it to any other alternative, i.e. if $\exists$ unique $j \in [n]$, s.t. $\forall s \in [n] : |\{i \in [m] : p_{ij} > p_{is}\}| > |\{i \in [m] : p_{ij} \leq p_{is}\}|$, then $p_j = 1$.

B. **Borda rule**: an alternative is chosen with certainty if it is the alternative with the highest probability using the average rule, i.e. let p_j defined as in equation (2) for all $j \in [n]$, and $J := \{\arg\max_{j \in [n]} p_j\}$. If $J = \{s\}$ for some s, then $p_s = 1$.

C. **Pareto rule**: an alternative is chosen with certainty if it dominates all other alternatives in the sense of Pareto, where alternative A_j dominates alternative A_k in the sense of Pareto (written $A_j \underset{pareto}{\succ} A_s$) iff. $A_j \succeq_i$ for all i and $A_j \succ_{i_0} A_k$ for some i_0. Pareto rule states that if $\exists j \in [n]$, s.t. $\forall s \in [n], A_j \underset{pareto}{\succ} A_s$, then $p_j = 1$.

D. **Intersection rule**: an alternative is chosen with certainty if it is the first choice of all individuals, i.e. $\forall i \in [m], J_i := \{\arg\max_{j\in[n]} q_{ij}\}$. If $\exists s \in [n]$, s.t. $\forall J_i = \{s\}$ for all i, then $p_s = 1$.

Above four rules do not satisfy condition (i) , i.e. they are not well-defined on P^m in the sense that to some probability matrix there does not exist corresponding social probability vector.

However, they can be modified to meet this condition, for instance:

A'. **Majority rule (modified)**: all alternatives in the set of alternatives which a majority prefers to all others receive equal social probability; if this set is empty, every alternative receives equal social probability[5]. Formally, let $J := \{j \in [n] : \forall s \in [n] : |\{i \in [m] : p_{ij} > p_{is}\}| > |\{i \in [m] : p_{ij} \leq p_{is}\}\}|$, if $|J| > 0,\ p_j = \frac{1}{|J|},\ \forall j \in J$; otherwise $p_j = \frac{1}{n},\ \forall j \in [n]$.

B'. **Borda rule (modified)**: all alternatives tied for the maximum probability using the average rule receive equal social probability. Using the notations in A, $p_s = \frac{1}{|J|},\ \forall s \in J$.

C'. **Pareto rule (modified)**: all alternatives not dominated in the sense of Pareto by some other alternative receive equal social probabilities, i.e. let $J := \{j \in [n] : \nexists s \in [n] \text{ s.t. } A_s \underset{pareto}{\succ} A_j\}$, then $p_j = \frac{1}{|J|},\ \forall j \in J$.

D'. **Intersection rule (modified)**: all alternatives in the lowest numbered non-empty social choice set receive equal social probability, i.e. let $J_i^{(r)}$ be the first r favorite alternatives of individual i, and the r-th social choice set $C^{(r)} := \bigcap_{i\in[m]} J_i^{(r)}$, let $s := \arg\min_{r\in[n]}(C^{(r)} \neq \emptyset)$, then $p_j = \frac{1}{|C^{(s)}|}, \forall j \in C^{(s)}$. This rule is an instance of the concept of a "dark horse candidate".

Quiz: Consider following a probability matrix:

$$\mathbf{Q} = \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.6 & 0.4 & 0 \\ 0 & 0.4 & 0.6 \end{pmatrix}$$

What are the social probabilities yield from the average rule and the modified four rules mentioned above?

4 Comparisons with rival rules

Compared to its rival rules, the average rule has following advantages and shortcomings:

[5] The empty-set case is not considered by the author in the original paper.

(a) Sensitivity to the strength of individual preference.[6] Using average rule, there are examples when individual preference ordering is the same, but the social preference is changed when one individual change her preference intensity. The other rules might not respond to such change, or respond discontinuously. If all the individuals are indifferent among all alternatives except one individual prefer one alternative with certainty, then as the number of individual grows larger and larger, the social preference would converge to social indifference.

(b) None of the rival rules satisfy the *unanimity preservation* condition. *Unanimity preservation* states that if all individuals choose an alternative with the same probability, then so does the society. This condition is always satisfied by average rule, but not by any of the rival rules.

(c) Symmetric matrix gives indifference of social preference: if a probability matrix is symmetric, then the social probability vector is uniform by the average rule.

(d) In terms of the probability that all individuals obtain their desired outcome [7], Borda rule exceeds the average rule in case of individuals with homogeneous probability vectors. It is a well-known result in individual choice theory that "probability matching" is less desirable than choosing the most preferred alternative with certainty. In this sense, Borda rule is preferable to the average rule in the case of one individual, and of $m \geq 2$ individuals with homogeneous probability vectors (but the result does not carry to the individuals with inhomogeneous probabilities[8]).

---

[6] The author suggests that the individual probabilities could be alternatively derived by giving each individual a certain number of votes and allowing her to distribute them among the alternatives in a manner proportional to her strength of preferences. However, Fishburn pointed out in [1] that the preference intensity elicited in the vote distribution method might be a different notion of the preference intensity interpreted from the choice probabilities. For example, for three alternatives, one may have choice probability vector $(1, 0, 0)$, and the distribution of 100 votes $(75, 25, 0)$; in addition, as a consequence, no social rule based solely on the individual's choice probabilities can distinguish among the individuals' second most preferred alternatives and their third most preferred alternatives and so forth.

[7] One can consider this to be one of the interpretations of the *degree to which the alternative chosen by the social choice rule agrees with the potential individual choices*, suggested by Fishburn in [1]. Another interpretation discussed by Fishburn is the *expected number of individual choices that coincide with the social choice*. In this sense, Fishburn gave a proof that Borda rule is always preferable than the average rule, in fact, it is the rule with the highest expected value.

[8] The result is uncertain for multiple individuals with different probabilities. The reader can consider a case of two individuals with $q_1 = (1, 0)$, $q_2 = (1/4, 3/4)$, and another case of two individuals with $q_1 = (0.1, 0.24, 0.25, 0.25, 0.16)$, $q_2 = (0.4, 0.25, 0.2, 0.15, 0)$. Borda rule exceeds the average rule in the first case, fails in the second case. One important remark is that, in the original paper, the author tries to use the first example to show that average rule is preferable than Borda rule, however, the calculation in his paper is problematic, as pointed out by Fishburn in [1].

Quiz: The result does not carry to inhomogeneous cases. Can you explain why? (Hint: use the examples given in the footnote.)

5 Some further properties of the average rule

In addition to axioms (i) - (iv) in section 2, the average rule satisfies a number of other important conditions:

(v) *Collective rationality.* Given any set of individual preferences, the implied social preferences form a complete preordering, i.e. it satisfies
completeness: $\forall j, k \in [n] : (A_j \succeq A_k) \vee (A_k \succeq A_j)$,
reflexivity: $\forall j \in [n] : A_j \succeq A_j$, and
transitivity: $\forall j, k, l \in [n] : A_j \succeq A_k, A_k \succeq A_l \implies A_j \succeq A_l$.
For any rule that outputs social probabilities and defines social preferences based on the value comparison of the probabilities, this condition always holds.

(vi) *Citizen's sovereignty.* Given any social probability vector, there exists a set of individual probability vectors which by the proposed rule, yields the given social probability vector. In other words, the social probability function f is surjective.

(vii) *Non-dictatorship.* No single individual or group of individuals smaller than the society itself can determine social probabilities independently of the other individuals, i.e. for fixed n alternatives, and a set of rules $\{f^{(k,n)} : k \in [m]\}$: $\forall I$ satisfying $\emptyset \neq I \subsetneq [m]$, define $k := |I|$, and $U(Q) := Q_{|I} = \{(q_i)_{i\in I} : Q = (q_i)_{i\in[m]}\}$ with $Q \in P^m$ (one obtains $U(Q)$ by choosing from Q rows with index in I), it holds that $f^{(k,n)} \circ U \neq f^{(m,n)}$ on P^m.

(viii) *Symmetry.* All individuals count equally in determining social probabilities in that social probabilities are unchanged if the individuals are relabeled, i.e. $\forall Q \in P^m$, let σ be some permutation (row-wise) operator on Q, then $f(\sigma(Q)) = f(Q)$, for $\forall\sigma$.[9]

(ix) *Unanimous preference preserving.* If all individuals prefer one alternative to another then so does society, and if all individuals are indifferent between two alternatives then so is society, i.e. $\forall j, k \in [n]$:
$\forall i \in [m]\ A_j \succ_i A_k \iff A_j \succ A_k$; and $\forall i \in [m]\ A_j \sim_i A_k \iff A_j \sim A_k$.

(x) *Pareto optimally.* If A_j dominates A_k in the sense of Pareto, then the society prefers A_j to A_k, i.e. if $A_j \underset{pareto}{\succ} A_k$, then $A_j \succ A_k$.

[9]One can prove that equal sensitivity defined in axiom (iv) implies permutation invariance of individuals.

(xi) *Certainty principle.* Society chooses/rejects an alternative with certainty if and only if all individuals choose/reject it with certainty. In other words, $\forall j \in [n]$ and $\forall \rho \in \{0,1\}$: $\forall i \in [m]$ $q_{ij} = \rho \iff p_j = \rho$.

(xii) *Unanimity preservation.* If all individuals choose an alternative with the same probability, then so does society, i.e. for $\omega \in P$: if $q_i = \omega$ for $\forall i \in [m]$ then $p = \omega$.

References

[1] Peter C. Fishburn. A probabilistic model of social choice: Comment. *The Review of Economic Studies*, 42(2):297–301, 1975.

[2] Michael D. Intriligator. A probabilistic model of social choice. *The Review of Economic Studies*, 40(4):553–560, 1973.

[3] P. M. Rice. Comments on a probabilistic model of social choice. *The Review of Economic Studies*, 44(1):187–188, 1977.