Arrow's Impossibility Theorem

Renuka Ravindran

In this article a number of voting models are considered and the drawbacks of each are indicated. The aim is to develop a voting method, which is based on individual preferences and which finally represents the choice of society. Arrow's theorem addresses this problem.

Renuka Ravindran was on the faculty of the Department of Mathematics, Indian Institute of Science, Bangalore. She was also Dean, Science Faculty at IISc. She has been a visiting Professor at various universities in USA and Germany.

Introduction

Voting is a prerequisite for any democratic process. Every individual, who is a member of a group/society/country, has the right to express his/her opinion and this opinion must be taken into account, while arriving at a final choice. This process of coming to a final choice could be varied and the possible ways of doing it are numerous. One would aspire to have an ideal voting process, which is fair, free, takes account of individual voter preferences and is not dominated by a single individual or a privileged group.

In a model for voting, we have two basic sets:

1. a set V of individuals who vote, denoted by $i, j, \cdot$, and

2. a set X of possible alternatives, denoted by $x, y, z, \cdot$.

The alternatives are mutually exclusive, in that at most one can be elected and the set X is complete in that it includes all possible outcomes.

In the process of voting, individuals express their 'preference' between alternatives. Each preference is a relative comparison of some pair of alternatives (see *Box* 1). For any pair of alternatives, say x and y, each individual i has a preference of x over y, denoted by xR_iy or has a

Keywords
Methods of voting, collective choice rule, almost decisive.

Box 1.

Relations Given a set $S = \{a, b, c, \cdots\}$, a subset R of the set $S \times S$ of ordered pairs of elements $(a, b), a \in S, b \in S$ is called a relation. If $(a, b) \in R$, we denote it by aRb. For example, S may be the set of all children of one generation in a joint family. The relation R could stand for 'brother of'. aRb if a is the brother of b. An equivalence relation on S, denoted by $\sim$, is a relation which has the following properties:

1. Transitive: If aRb and bRc, then aRc.
2. Symmetric: If aRb, then bRa.
3. Reflexive: aRa for all S.

If S is the set of all triangles in the plane, R could be 'congruent to'.

preference of y over x, denoted by yR_ix or is indifferent between x and y. We assume that each individual's preferences are 'self-consistent'. Each individual who prefers x over y and y over z, also prefers x over z.

$$xR_iy \wedge yR_iz \rightarrow xR_iz.$$

The main question is the following:

Based on individual preferences R_i, how can a choice R be made, which represents the choice of society. The precise details of how the individuals or groups express their preferences is of no consequence and the details of how the collective choice is arrived at is of no importance. However, the choice must satisfy certain conditions of 'reasonableness'. Can one impose certain conditions on the choice? In his 'Impossibility theorem', Arrow proved that a set of very 'mild' conditions are so restrictive that they rule out not some, but *every possible*, 'reasonable' choice.

Let us first look at some examples of voting models.

2. Example 1: Method of Majority Decision

Suppose there are 100 members of a House of Representatives, who have the right to vote. Three issues x, y and z are to be put up for voting and the leader decides to use the method of majority decision – two issues are put up for vote at a time and the one with the larger number of votes wins. Of the 100 members:

33 prefer x to y and y to z,
35 prefer y to z and z to x,
32 prefer z to x and x to y.

There are no other preferences.

If the leader of the House first puts up y and z for vote, then 33+35 = 68 vote for y and 32 vote for z. y wins over z or yRz.

Now the leader puts up y and x for vote, then 33+32 = 67 vote for x and 33 vote for y. x wins over y or xRy. So the House votes x as the winner.

Had the leader first put up z and x for vote, then we would have had zRx. Next if z and y are put to vote, then yRz. So y is the winner.

By first putting x and y to vote, x would win and then between x and z, z would win.

It therefore emerges that the order in which the issues are put up for vote determines the choice of the House. This is far too arbitrary to be acceptable. Besides, the choice of the House by the method of majority decision is not transitive.

If y and z are put up for vote first, then yRz.
If x and y are put up for vote next, then xRy.
But if x and z are put up for vote, then zRx.
So $yRz \wedge xRy \not\rightarrow xRz$.

Transitivity is a property that the collective choice must have.

3. Example 2: Rank Order Method

Suppose that we consider now another mothod of voting, common at contests, for example, beauty contests or painting or dance competitions, where a certain subjective element is involved in the marking. In a rank order method of voting, each of the judges rates the candidates as first, second or third, etc. Marks are given, say, 5 for being first, 3 for being second, 1 for being third and 0 for the rest. Let us consider a situation, where there are four candidates w, x, y, z and four judges i, j, k, l. Suppose

	w	x	y	z
i	5	0	1	3
j	5	1	3	0
k	3	1	5	0
l	0	1	3	5
Total	13	3	12	8

Then w is declared the winner and y the runner up. Suppose on some grounds, z is disqualified, then the judges' rating between w and y remains unchanged. Now

	w	x	y
i	5	1	3
j	5	1	3
k	3	1	5
l	1	3	5
Total	14	6	16

Now y becomes the winner and w the runner up. Even though all the judges have not changed the relative rating between w and y, the collective rating has changed because of the disqualification of z. This method of voting has the defect that the collective choice depends on the presence or absence of another alternative. Ideally, a voting method should be independent of an irrelevant alternative.

4. Example 3: Traditional Code Method

A choice based on some traditional code – eldest son or seniority. Then, even if all the voting members prefer x to y, the code may prefer y to x, for example. This

Box 2.

Kenneth J Arrow was born in the City of New York on August 23, 1921. He graduated in 1940 from the City College in New York with a Bachelor of Science in social science but a major in mathematics, a paradoxical combination that was in tune with his future interests. He is best known for his 1951 PhD dissertation (on which his book *Social Choice and Individual Values* was based. In 1972, Arrow, jointly with Sir John Hicks, won the Nobel Prize in Economics. It was awarded for his 'pioneering contributions to general equilibrium theory and welfare theory.' Arrow has spent most of his professional life on the Economic faculties of Stanford University and Harvard University.

defies Pareto's principle – if everyone prefers x to y, then the collective choice must also prefer x to y.

Hence methods of combining individual preferences into a collective social proference could lead to inconsistencies.

5. Arrow's Theorem

Arrow's (see *Box* 2) Impossibility Theorem consists of imposing certain conditions on a collective choice rule.

Condition U (Unrestricted Domain): The domain of the rule must include all possible combinations of individual preferences.

Condition P (Pareto's Principle): For any pair x, y in X, if every individual prefers x to y, then the collective choice rule must also prefer x to y.

$$[\ \forall i : xR_i y\ ] \rightarrow xRy.$$

Condition I (Independence of Irrelevant Alternative): Let S and S' be two subsets of X and suppose $(R_1, R_2,$ $, R_n)$ and $(R'_1, R'_2, \cdot \quad , R'_n)$ are two sets of individual preferences on S and S', respectively. If a pair of alternatives, x and y are in both S and S' and if $xR_i y \rightarrow xR'_i y$ for all i, then the collective choice between x and y must also be the same in both cases. It must be independent of the presence or absence of a third alternative z.

Condition D (Nondictatorship): There is no individual i, such that for every pair of alternatives x, y in $X, xR_iY \to xRy$.

Theorem 1. (Arrow's Theorem) There is no collective choice function satisfying U, P, I and D.

Definition 1: A set of individuals V is *almost decisive* for x against y if xRy, whenever xR_iy for every i in V and yR_jx for every j not in V This is denoted by $V : D(x, y)$.

Definition 2: A set of individuals V is *decisive* for x against y if xRy, whenever xR_iy for every i in V This is denoted by $V : \tilde{D}(x, y)$.

According to definition, a person is *almost decisive* if he wins when there is opposition and he is *decisive* if he wins whether he is opposed or not. It is natural to assume that if a person wins in spite of opposition, he must win when others do not oppose him. However, such an assumption is not made. In the absence of it, to be decisive is stronger than being almost decisive, because the former implies the latter, but not vice versa. $V : \tilde{D}(x, y) \to V : D(x, y)$.

Lemma 1. *If there is an individual J, who is almost decisive for any ordered pair of alternatives, then a collective choice rule satisfying U, P and I implies that J must be a dictator.*

Proof: J is almost decisive for a given pair of alternatives, that is, $\exists\ x, y \in X, \ni J : D(x, y)$. Let z be any other alternative and let i refer to all individuals other than J.

Suppose xR_Jy and yR_Jz and suppose further yR_ix and yR_iz. We conclude xR_Jz since an ordering is a transitive relation. Only J has specified a preference between x and z. None of the others has expressed a preference between x and z.

Since $J : D(x, y)$ and $[xR_Jy \wedge yR_ix]$, it follows that xRy. Also $[yR_Jz \wedge yR_iz] \rightarrow yRz$ by the unanimity condition. So we have $xRy \wedge yRz$. This implies xRz. This result is arrived at without any assumption about the preference regarding x and z other than by J. Of course, we have assumed yR_iz and yR_ix. If these rankings of y vis-a-vis x and y vis-a-vis z have any effect on the collective choice between x and z, then we violate condition I (independence of irrelevant alternative). Hence xRz must be independent of these assumptions. That means $J : D(x, y) \rightarrow J : \tilde{D}(x, z)$.

Now suppose zR_jx and xR_Jy, while zR_ix and yR_ix. The only preference between z and y is expressed by J alone and no one else. $J : D(x, y)$ and $[xR_Jy \wedge yR_ix] \rightarrow xRy$. By unanimity, $[zR_Jx \wedge zR_ix] \rightarrow zRx$. Now $xRy \wedge zRx \rightarrow zRy$ and this with only zR_Jy, without anything being specified about the preference of other individuals between y and z. $J : D(x, y) \rightarrow J : \tilde{D}(z, y)$.

We now have

$$J : D(x, y) \rightarrow J : \tilde{D}(x, z) \tag{1}$$

and

$$J : D(x, y) \rightarrow J : \tilde{D}(z, y). \tag{2}$$

Interchanging y and z in (2), we get

$$J : D(x, z) \rightarrow J : \tilde{D}(y, z). \tag{3}$$

By a cyclic change $x \rightarrow y, y \rightarrow z, z \rightarrow x$ in (1), we have

$$J : D(y, z) \rightarrow J : \tilde{D}(y, x). \tag{4}$$

We now have

$$\begin{aligned} J : D(x, y) &\rightarrow J : \tilde{D}(x, z) \\ &\rightarrow J : D(x, z) \quad \text{by definition} \end{aligned}$$

$$\begin{array}{ll} \rightarrow \; J : \tilde{D}(y,z) & \text{from (3)} \\ \rightarrow \; J : D(y,z) & \text{by definition} \\ \rightarrow \; J : \tilde{D}(y,z) & \text{by (4)} \end{array}$$

$$J : D(x,y) \rightarrow \{J : \tilde{D}(x,z) \wedge J : \tilde{D}(z,y) \wedge J : \tilde{D}(y,x)\}.$$

Interchanging x and y

$$J : D(y,x) \rightarrow \{J : \tilde{D}(y,z) \wedge J : \tilde{D}(z,x) \wedge J : \tilde{D}(x,y)\}.$$

But $J : D(x,y) \rightarrow J : \tilde{D}(y,x) \rightarrow J : D(y,x)$.

So $J : D(x,y) \rightarrow \{J : \tilde{D}(x,z) \wedge J : \tilde{D}(z,y) \wedge J : \tilde{D}(y,x) \wedge J : \tilde{D}(y,z) \wedge J : \tilde{D}(z,x) \wedge J : \tilde{D}(x,y)\}$.

J is decisive for every ordered pair of alternatives from the set of three alternatives (x,y,z), provided $J : D(x,y)$ and given conditions U, P and I. So J is a dictator over any set of three alternatives containing x and y.

Now consider a large number of alternatives. Take any two alternatives u and v out of this set.

If u and v are chosen as x and y, then $J : D(x,y) \rightarrow \{J : \tilde{D}(u,v) \wedge J : \tilde{D}(v,u)\}$ by taking a triple consisting of u, v and any other alternative z and using the result proved earlier.

If u is the same as x, but v is diferent from y, then considering the triple (u,v,y), we have $J : D(x,y) \rightarrow \{J : \tilde{D}(u,v) \wedge J : \tilde{D}(v,u)\}$.

If u and v are both different from x and y, first consider the triple (x,y,u). Then $J : D(x,y) \rightarrow \{J : \tilde{D}(u,x) \wedge J : \tilde{D}(x,u)\}$. Now $J : \tilde{D}(x,u) \rightarrow J : D(x,u)$ and considering the triple (x,u,v), we have $J : D(x,u) \rightarrow \{J : \tilde{D}(u,v) \wedge J : \tilde{D}(v,u)\}$. Thus $J : D(x,y)$ for some x,y implies $J : \tilde{D}(u,v)$ for all possible ordered pairs (u,v). Therefore individual J is a dictator.

Proof of the Theorem: For every pair of alternatives, there is at least one decisive set, namely, the set of all individuals, thanks to condition P. For every pair of alternatives, there is also at least one almost decisive set (since every decisive set is almost decisive). For each pairwise choice of alternatives, consider the smallest one among them (or one of the smallest ones). Let this set V be almost decisive for the ordered pair of alternatives x and y.

If V contains one individual, then the theorem is proved because of the lemma.

If V contains more than one individual, then divide V into two parts. V_1 contains a single individual and V_2 contains the rest of V All indiviuals not in V form the set V_3, which could be empty. Due to the condition U, we can choose any logical possible combination for individual preferences. We choose the following for x, y and a third alternative z:

1. for all i in V_1, $xR_iy \wedge yR_iz$,
2. for all j in $V_2, xR_jy \wedge zR_jx$,
3. for all k in $V_3, yR_kz \wedge zR_kx$.

We are given $V : D(x, y)$ and $[xR_iy \wedge xR_jy]$ for $i, j \in V$ and yR_kz for all k not in V, so xRy. Between y and z, only V_2 prefers z to y, while V_1 and V_3 prefer y to z. If zRy, then $V_2 : D(z, y)$, but this is not possible, since V is the smallest almost decisive set. So zRy does not hold. For R to be total as needed for condition U, yRz must hold. $[xRy \wedge yRz] \to xRz$. Only V_1 prefers x to z, so that a certain individual has turned out to be almost decisive. This contradicts the assumption.

An individual almost decisive over some pair must be a complete dictator.

Suggested Reading

[1] **Arrow, Kenneth J,** ***Social Choice and Individual Values*****, John Wiley NY, 1951.**

Address for Correspondence
Renuka Ravindran
Department of Mathematics
Indian Institute of Science
Bangalore 560 012, India.